

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 13:30:51 ; Search time 20.3248 Seconds
(without alignments)
932.422 Million cell updates/sec

Title: US-09-492-971B-15_COPY_4_52
Perfect score: 275
Sequence: 1 QAQMVPQSPVAVSQSRPG.....INQWERTYLGVLVCTCYG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	275	100.0	259	AA90281	Aay90281 Human fib
2	275	100.0	720	AA28914	Aay28914 Fibronect
3	275	100.0	2324	AA922778	Aar922778 Human fib
4	275	100.0	2324	AAU74674	Aau74674 Human fib
5	275	100.0	2324	AAE23651	Aae23651 Human pro
6	275	100.0	2327	AA70373	Aap70373 Human fib
7	275	100.0	2327	AAE15468	Aar15468 Human fib
8	275	100.0	2328	AA68182	Aag68182 Fibronect
9	275	100.0	2328	ABU07486	Abu07486 Protein d
10	275	100.0	2328	ABR41106	Abr41106 Human fib
11	275	100.0	2328	ABR92078	Abr92078 Human cer
12	275	100.0	2328	AD70378	Ad70378 Fibronect
13	275	100.0	2328	ADB98726	Adb98726 Human fib
14	275	100.0	2328	ADE82522	Ad82522 Human pro
15	275	100.0	2328	ADJ37157	Adj37157 Human mal
16	275	100.0	2386	AAW63171	Aaw63171 Amino aci
17	275	100.0	2446	AAE60021	Aae60021 Fibrinoge
18	275	100.0	2474	ABG22279	Abg22279 Novel hum
19	275	100.0	2477	AAW99595	Aaw99595 Human fib
20	272	98.9	1179	ADP75952	Adp75952 Human min
21	272	98.9	1359	ADP75957	Adp75957 Human leu
22	271	98.5	231	AA58210	Aab58210 Lung canc
23	271	98.5	463	ABR58303	ABr58303 BCU0770 p
24	271	98.5	642	ADQ39403	Adq39403 Human myo
25	271	98.5	642	ADR67316	Adr67316 Human bia

ALIGNMENTS

RESULT 1
AA90281
ID AA90281 standard; protein; 259 AA.

XX AC AA90281;
XX DT 13-OCT-2000 (first entry)
XX DE Human fibronectin protein sequence fragment.

KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; fibronectin.

XX OS Homo sapiens.

XX PN EP1024192-A2.

XX PD 02-AUG-2000.

XX PF 23-DEC-1999; 99EP-00310541.

XX PR 24-DEC-1998; 98IN-DE003825.

XX PA (COUL) CSIR COUNCIL SCI IND RES.

XX PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
XX PI Yadav M;

XX WPI: 2000-516032/47.

XX DR N-PSDB; AAA37632.

XX PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX PT domains of human fibronectin.

XX Example 3; Fig 6; 58pp; English.

XX PS This sequence represents a human fibronectin fragment, containing fibrin
XX PS binding domains. The invention relates to a hybrid plasminogen activator
XX PS (PA) comprises a polypeptide fusion between streptokinase (SK), which are
XX PS capable of plasminogen (PG) activation, and fibrin binding regions of
XX PS human fibronectin, which are from fibrin binding domains (FBD) 4 and 5 or
XX PS 1 and 2. The hybrid PA possesses the ability to bind with fibrin
XX PS independently and also characteristically retains a PG activation ability
XX PS which becomes evident only after a pronounced duration, or lag, after
XX PS exposure of the PA to a suitable animal or human PG. The hybrid
XX PS streptokinase-fibrin binding domain polypeptides are useful in

26	271	98.5	642	8	ADS17489	Adsl7489 Amino aci
27	271	98.5	642	8	ADR97658	Adr97658 Human fib
28	271	98.5	657	8	ADQ39409	Adq39409 Human myo
29	271	98.5	660	2	AA928901	Aay28901 Human mig
30	271	98.5	984	8	ADQ39406	Adq39406 Human myo
31	271	98.5	1173	4	ABG22275	Abg22275 Novel hum
32	271	98.5	2182	8	ADR66462	Adr66462 Human pro
33	271	98.5	2182	8	ADR66120	Adr66120 Human pro
34	271	98.5	2220	6	ABO01289	Ab001289 Human pro
35	271	98.5	2220	8	ADN95950	Adn95950 Human NOV
36	271	98.5	2265	4	AA38647	Aam38647 Human cel
37	271	98.5	2266	6	ABR40124	ABr40124 Human cel
38	271	98.5	2296	8	ADQ39412	Adq39412 Human myo
39	271	98.5	2330	4	AA38646	Aam38646 Human pol
40	271	98.5	2355	4	AA38649	Aam38649 Human pol
41	271	98.5	2355	6	ABR58335	ABr58335 NW_00202
42	271	98.5	2355	7	ADP65196	Adp65196 Human fib
43	271	98.5	2355	8	ADG99560	Adg99560 Human fib
44	271	98.5	2355	8	ADL92160	Adl92160 Fibronect
45	271	98.5	2355	8	ADO55175	Ado55175 Protein #

CC thrombolytic therapy for various kinds of cardiovascular disorders. The
CC hybrids have enhanced fibrin selectivity as well as kinetics of
CC plasminogen activation that are distinct from that of natural
CC streptokinase in being characterised by a temporary delay, or lag of
CC several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase
XX
XX
SQ Sequence 259 AA;

Query Match 100.0%; Score 275; DB 3; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.8e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQMVQPSPVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYG 49
|||||
DB 1 QAAQMVQPSPVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYG 49
|||||

RESULT 2
AAY28914
ID AAY28914 standard; protein; 720 AA.

XX AAY28914;

XX 21-SEP-1999 (first entry)

XX Fibrinectin protein sequence.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibrinectin.

XX Homo sapiens.

XX WO9931233-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.

XX Disclosure; Fig 2; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. The present sequence
CC represents the human fibrinectin

XX Sequence 720 AA;

Query Match 100.0%; Score 275; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 6e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQMVQPSPVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYG 49
|||||

Db 50 QAAQMVQPSPVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYG 98
RESULT 3
AAR92778
ID AAR92778 standard; protein; 2324 AA.
XX AAR92778;
XX 21-OCT-2004 (revised)
DT 21-JUN-1996 (first entry)
XX Human fibrinectin.
XX Fibrin-binding peptide; fibrinectin; fibrinolysis; fibrinogenesis;
KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis; therapy.
XX Homo sapiens.
OS Unidentified.
XX Key Location/Qualifiers
FT Domain 21..65
/label = 1F1
/note = "type 1 module 1"
FT Domain 66..109
/label = 2F1
/note = "type 1 module 2"
FT Domain 110..154
/label = 3F1
/note = "type 1 module 3"
FT Domain 155..199
/label = 4F1
/note = "type 1 module 4"
FT Domain 200..244
/label = 5F1
/note = "type 1 module 5"
FT Domain 277..312
/label = 6F1
/note = "type 1 module 6"
FT Domain 439..486
/label = 7F1
/note = "type 1 module 7"
FT Disulfide-bond 439..467
/note = "Cys439-Cys467 disulfide bond"
FT Disulfide-bond 465..477
/note = "Cys465-Cys477 disulfide bond"
FT Domain 487..529
/label = 8F1
/note = "type 1 module 8"
FT Domain 530..569
/label = 9F1
/note = "type 1 module 9"
FT Domain 2123..2230
/label = Fibrin binding domain
/note = "11 kDa C-terminal fibrin-binding domain"
FT Domain 2141..2185
/label = 10F1
/note = "type 1 module 10"
FT Disulfide-bond 2144..2173
/note = "Cys2144-Cys2173 disulfide bond"
FT Disulfide-bond 2171..2183
/note = "Cys2171-Cys2183 disulfide bond"
FT Domain 2187..2230
/label = 11F1
/note = "type 1 module 11"
FT Disulfide-bond 2189..2216
/note = "Cys2189-Cys2216 disulfide bond"
FT Disulfide-bond 2214..2226
/note = "Cys2214-Cys2226 disulfide bond"
FT Domain 2233..2271
/label = 12F1
/note = "type 1 module 12"
XX

PN WO9604304-A1.
 XX 15-FEB-1996.
 PD
 XX 01-AUG-1995; 95WO-US009819.
 PF
 XX 01-AUG-1994; 94US-00283857.
 PR
 XX (UUNY) UNIV NEW YORK STATE.
 PA
 XX Gold LI, Rostagno AA;
 PI WPI; 1996-129333/13.
 XX DR N-PSDB; AAT17551.
 XX DR
 XX New fibrin-binding peptide molecules - used for the diagnosis and
 PT treatment of conditions associated with fibrin deposition, e.g. thrombi.
 PT
 XX Claim 2; Page 104-110; 146pp; English.
 PS
 XX Human fibronectin (AAR92778) has a fibrin-binding site, close to the C-
 CC terminus (amino acids 2123-2232 or 2141-2230), covering the 10F1.11F1
 CC module pair of the protein. Fibrin-binding sites can be prep'd. by
 CC enzymatic cleavage of fibronectin, peptide synthesis or by recombinant
 CC DNA techniques. They are used to detect a fibrin-binding target site, to
 CC treat disorders involving abnormal fibrinolysis or fibrinogenesis, to
 CC disrupt blood clots and to aid delivery of medicaments to fibrin- contg.
 CC sites. They show high affinity to and slow dissociation from fibrin, and
 CC provide fast diffusion and rapid clearance
 CC
 CC Revised record issued on 21-OCT-2004 : Correction to Feature Table Key
 CC
 XX Sequence 2324 AA;
 SQ

Query Match 100.0%; Score 275; DB 2; Length 2324;
 Best Local Similarity 100.0%; Pred. No. 2.4e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQWVQSPVAVSQSGKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
 |||||
 DB 1 QAAQWVQSPVAVSQSGKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
 |||||

RESULT 4
 AAU74674
 ID AAU74674 standard; protein; 2324 AA.
 XX
 XX AAU74674;
 AC
 XX
 XX 09-APR-2002 (first entry)
 DT
 XX Human fibronectin protein.
 DE
 XX Human; fibronectin; VLA-4 binding site; very late antigen-4;
 KW quiescent haematopoietic cell; HC; apoptosis; CD34+.
 KW
 XX Homo sapiens.
 OS
 XX WO200187071-A1.
 PN
 XX 22-NOV-2001.
 PD
 XX 12-MAY-2000; 2000WO-US012993.
 PF
 XX 12-MAY-2000; 2000WO-US012993.
 PR
 XX (ADRE-) ADVANCED RES & TECHNOLOGY INST.
 PA
 XX Williams DA, Bradford GB, Dutt P, Yoder MC;
 PI WPI; 2002-082932/11.
 XX
 XX Obtaining hematopoietic cell population containing quiescent cells for

PT use in treating a subject, by expanding the cells while adhered to a
 PT polypeptide containing binding sites for integrins on the cells.
 XX
 PS Disclosure; Page 51-61; 63pp; English.
 XX
 XX The invention relates to obtaining a population of quiescent
 CC haematopoietic cells (HC), comprising culturing HC while adhering the
 CC cells to a polypeptide having a very late antigen-4 (VLA-4) binding site,
 CC in particular fibronectin polypeptide, so as to expand the number of HC,
 CC where the adhering provides an increased percentage of quiescent HC. Also
 CC included are inducing apoptosis of a subpopulation of HC, by contacting
 CC the cells with a polypeptide having a VLA-4 binding site under conditions
 CC to cause apoptosis of a subpopulation of HC and a medium for culturing HC
 CC which enriches quiescent HC, comprising a fibronectin polypeptide. The
 CC method is useful for obtaining a cell population containing quiescent
 CC haematopoietic cells which are useful for inducing apoptosis of a
 CC VLA-4 containing peptide is useful for inducing apoptosis of a
 CC subpopulation of CD34+ HC population. The present sequence is human
 CC fibronectin which contains then VLA-4 binding site used in the method of
 CC the invention
 XX
 SQ Sequence 2324 AA;
 Query Match 100.0%; Score 275; DB 5; Length 2324;
 Best Local Similarity 100.0%; Pred. No. 2.4e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQWVQSPVAVSQSGKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
 |||||
 DB 1 QAAQWVQSPVAVSQSGKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
 |||||

RESULT 5
 AAE23651
 ID AAE23651 standard; protein; 2324 AA.
 XX
 XX AAE23651;
 AC
 XX 10-SEP-2002 (first entry)
 DT
 XX Human protein related to modulation of proliferation of EPCs.
 DE
 XX Human; proliferation; apoptosis; erythroid progenitor cell; EPC; VLA;
 KW integrin very late antigen; haematopoietic cell; induced disorder;
 KW genetic disorder; congenital erythroid disorder; polycythaemia; anaemia.
 KW
 XX Homo sapiens.
 OS
 XX WO200187037-A1.
 PN
 XX 22-NOV-2001.
 PD
 XX 12-MAY-2000; 2000WO-US012961.
 PF
 XX 12-MAY-2000; 2000WO-US012961.
 PR
 XX (ADRE-) ADVANCED RES & TECHNOLOGY INST.
 PA
 XX Williams DA, Kapur R, Cooper R, Zhang L;
 PI WPI; 2002-121963/16.
 XX
 XX Modulating proliferation of erythroid progenitor cells in a hematopoietic
 PT cellular population by culturing with polypeptides containing an integrin
 PT very late antigen 5 or 4 binding site, useful in treating erythroid
 PT disorders.
 PT
 XX Disclosure; Page 55-64; 66pp; English.
 PS
 XX The present invention relates to a method for enhancing proliferation or
 CC inducing apoptosis of erythroid progenitor cells (EPCs). The method
 CC involves culturing haematopoietic cellular populations containing EPCs in
 CC the presence of a polypeptide having an integrin very late antigen (VLA)-

CC 5 or VLA-4 binding site. The EPCs are useful for treating patients having
CC induced, genetic or congenital erythroid disorders. Methods of the
CC invention are useful in diagnostic assays of progenitor cells, e.g. cells
CC occurring in pathologic states such as anaemia or polycythaemia and in
CC the study of erythroid cell development. They are used in the screening
CC of agents which inhibit or prevent apoptosis of erythroid cells. The
CC present sequence is a human protein related to the modulation of
CC proliferation of EPCs. This sequence is used in the invention
XX
XX
SQ Sequence 2324 AA;

Query Match 100.0%; Score 275; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQSPVAVSQSPGCGYDNGKHQYQINQOWERTYLGNVLVCTCYG 49
Dy 1 QAOQWVQSPVAVSQSPGCGYDNGKHQYQINQOWERTYLGNVLVCTCYG 49

RESULT 6
AAP70373
ID AAP70373 standard; protein; 2327 AA.

XX AAP70373;
XX
XX
DT 25-MAR-2003 (revised)
DT 11-MAR-1991 (first entry)

DE Human fibronectin gene product.

XX FN; collagen; fibrin; heparin.

XX Homo sapiens.

XX EP207751-A.

XX 07-JAN-1987.

XX 27-JUN-1986; 86EP-00304998.

XX 28-JUN-1985; 85GB-00016421.

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

XX Baralle FE;

XX WPI; 1987-001441/01.

DR N-PSDB; AAN70596.

XX New fibronectin polypeptide sequence with affinity for collagen etc. -
PT useful for targeting therapeutic substances on natural fibrin, for use
PT in affinity purificn. of polypeptide(s) etc.

PS Claim 11; Fig 3A; 32pp; English.

XX The product may be expressed from a transformed micro-organism, esp.
CC E.coli. FN binds to fibrin, heparin and Staphylococcus aureus, and may be
CC used to target a therapeutic agent onto natural fibrin eg. a blood clot.
CC It may also be used in affinity purification of a polypeptide, conjugated
CC to the collagen binding site of FN and immobilised on a collagen surface.
CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 2327 AA;

Query Match 100.0%; Score 275; DB 1; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQSPVAVSQSPGCGYDNGKHQYQINQOWERTYLGNVLVCTCYG 49
Dy 4 QAOQWVQSPVAVSQSPGCGYDNGKHQYQINQOWERTYLGNVLVCTCYG 52

RESULT 7
AAR15468
ID AAR15468 standard; protein; 2327 AA.

XX AAR15468;

DT 25-MAR-2003 (revised)

DT 12-MAR-1992 (first entry)

XX Human fibronectin.

DE Fibrin-imaging; atherosclerosis; thrombus inhibitor.

XX Homo sapiens.

XX WO9117765-A.

XX 28-NOV-1991.

XX 21-MAY-1990; 90US-00526397.

XX 21-MAY-1990; 90US-00526397.

XX (BIOT-) BIO-TECHN GEN CORP.

XX Vogei T, Levanon A, Werber M, Guy R, Panet A, Hartman J;

PI Shaked H;

XX WPI; 1991-369004/50.

XX N-PSDB; AAQ15214.

XX New fibrin binding domain polypeptide(s) - useful in imaging fibrin-

PT contg. substances, to inhibit thrombus formation and treat wounds.

XX Disclosure; Fig 2; 191pp; English.

XX The amino acid sequence is that of human fibronectin, this can be used to
CC derive polypeptides which are identical to part of the fibrin- binding
CC domain (FBD) of fibronectin. These polypeptides can be used to inhibit
CC thrombus formation; or (coupled to a thrombolytic agent) to induce
CC thrombolysis, or to treat wounds, e.g. in skin, eyes or tendons (in
CC conjunction with a polypeptide which includes a part of the cell-binding
CC domain (CBD) of fibronectin). These polypeptides are easier to prepare
CC than the full 31kD polypeptide. It can also be used to image fibrin-
CC contg. materials, esp. a thrombus or athero- sclerotic plaque, pref.
CC using a gamma counter. (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 2327 AA;

Query Match 100.0%; Score 275; DB 2; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQSPVAVSQSPGCGYDNGKHQYQINQOWERTYLGNVLVCTCYG 49

Dy 4 QAOQWVQSPVAVSQSPGCGYDNGKHQYQINQOWERTYLGNVLVCTCYG 52

RESULT 8

AAG68182

ID AAG68182 standard; protein; 2328 AA.

XX AAG68182;

DT 25-JAN-2002 (first entry)

XX Fibronectin protein SEQ ID NO:98.

XX Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
KW sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
KW antisense therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
KW osteomalacia; fibrous dysplasia.

XX Homo sapiens.
OS WO200177327-A1.
XX 18-OCT-2001.
XX 21-JUN-2000; 2000WO-US016951.
XX 05-APR-2000; 2000US-00543771.
XX 05-APR-2000; 2000US-00544398.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX WPI; 2001-657171/75.
XX New high bone mass (HBM) and Zmax1 genes and proteins useful for
XX modulating bone mass for the treatment of e.g. osteoporosis.
XX Claim 76; Page 408-413; 443pp; English.
XX The present invention describes the human Zmax1 gene and the high bone
XX mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
XX genes have osteopathic activities. The genes can be used in gene therapy,
XX antisense therapy and in the production of vaccines. They can be used in
XX the diagnosis and treatment of bone disorders including osteoporosis,
XX Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
XX ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
XX the exemplification of the present invention
XX
SQ Sequence 2328 AA;
Query Match 100.0%; Score 275; DB 4; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQWVQSPVAVSQKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
Db 5 QAAQWVQSPVAVSQKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 53
RESULT 9
ID ABU07486
AC ABU07486;
XX
DT 28-JAN-2003 (first entry)
DE Protein differentially regulated in prostate cancer #89.
XX
KW Prostate cancer; gene expression; differential regulation;
KW molecular marker; drug target; cancer detection; cancer diagnosis;
KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX
OS Homo sapiens.
XX WO200281638-A2.
XX
XX 17-OCT-2002.
XX
XX 08-APR-2002; 2002WO-US010824.
XX
XX 06-APR-2001; 2001US-0281731P.
XX 06-APR-2001; 2001US-0281732P.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX Sun Z, Jay G;
XX WPI; 2003-058520/05.
XX

DR N-PSDB; ABX10391.
XX
XX Novel genes which are differentially regulated in prostate cancer, useful
PT for diagnosing prostate cancer in prostate tissue sample and assessing
PT therapeutic or preventive intervention in prostate cancer patients.
XX
XX Claim 1; Page 378-386; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
XX prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
XX sample comprising prostate tissue, which involves determining the number
XX of target genes which are differentially-regulated in the sample, where
XX the number is indicative of the probability that the sample comprises
XX prostate cancer. (I) Is useful for assessing a therapeutic or preventive
XX intervention in a subject having a prostate cancer, which involves
XX determining the expression levels in a sample comprising prostate tissue
XX of target genes which are differentially-regulated in prostate cancer.
XX Preferably, the expression levels of at least 10 genes are determined.
XX (I) is also useful for identifying agents that modulate a biological
XX activity of a polypeptide differentially-regulated in prostate cancer
XX cells, which involves contacting a polypeptide differentially-regulated
XX in prostate cancer cells with a test agent under conditions effective for
XX the test agent to modulate a biological activity of the polypeptide, and
XX determining whether the test agent modulates the biological activity. (I)
XX is useful as molecular markers, as drug targets, and for detecting,
XX diagnosing, staging, grading, assessing, monitoring, prognosticating,
XX preventing or treating, determining predisposition to diseases and
XX conditions especially relating to prostate cancer. (I) and its expression
XX products are used in the diagnostic test to assay for presence of cancer
XX e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
XX blood etc. (I) is useful for assessing cancer e.g., to determine the type
XX of cancer, its stage of development, the nature of genetic defect, etc.
XX The polypeptide encoded by (I) can be used as target for therapy or drug
XX discovery. (I) can also be used for expressing the polypeptide and thus
XX for searching specific binding partners of the polypeptide. (I) is useful
XX in therapeutic applications to treat prostate cancer. The identification
XX of specific genes, and groups of genes, expressed in pathways
XX physiologically relevant to prostate cancer permits the definition of
XX functional and disease pathways and the delineation of targets in these
XX pathways which are useful in diagnostic, therapeutic, and clinical
XX applications. This is the amino acid sequence of a protein differentially
XX regulated in prostate cancer
XX
SQ Sequence 2328 AA;
Query Match 100.0%; Score 275; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQWVQSPVAVSQKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
Db 5 QAAQWVQSPVAVSQKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 53
RESULT 10
ID ABR41106
XX
XX ABR41106;
XX
XX 02-JUN-2003 (first entry)
XX
XX Human fibronectin gene FN protein product.
XX
XX Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone.
XX
XX Homo sapiens.
XX
XX WO200292764-A2.
XX

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
PA Gordon GJ, Jensen RV, Gullians SR, Bueno R;
XX WPI; 2003-290233/28.
DR N-PSDB; ADB70377.
XX
XX Diagnosing cancer cells in tissue sample, or determining prognosis or
PT outcome of cancer patient, by calculating ratio of expression levels of
PT genes that are differentially expressed in cancer and non cancer tissues.
XX
XX Claim 77; Page 369-376; 396pp; English.
XX
XX The present invention describes a method (M1) for diagnosing the presence
CC of cancer cells or non-cancer cells in a tissue sample, or determining
CC the prognosis or outcome of a cancer patient. M1 involves providing a set
CC of genes that are differentially expressed in cancerous or non-cancerous
CC conditions, determining the expression levels of the set of genes and
CC calculating a ratio of the expression levels of the differentially
CC expressed genes. M1 is useful for diagnosing the presence of cancer cells
CC or non-cancer cells in a tissue sample, where the cancer is malignant
CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma,
CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell
CC lymphoma, follicular lymphoma and ovarian cancer, and for determining
CC prognosis or outcome of a cancer patient. The ratio of expression levels
CC of differentially expressed genes is used as an indicator of cancer type,
CC cancer class, and/or cancer prognosis, all of which are useful for
CC determining a course of treatment of a patient. The present sequence
CC represents a human protein which is used in an example from the present
CC invention.

XX SQ Sequence 2328 AA;

Query Match 100.0%; Score 275; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
Db 5 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 53

RESULT 13

ADB98726
ID ADB98726 standard; protein; 2328 AA.

XX ADB98726;

XX 04-DEC-2003 (first entry)

XX Human fibronectin.

XX Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;
KW bone mass modulation; osteoporosis.

XX Homo sapiens.

XX WO200292000-A2.

XX 21-NOV-2002.

XX 13-MAY-2002; 2002WO-US014877.

XX 11-MAY-2001; 2001US-0290071P.

XX 17-MAY-2001; 2001US-0291311P.

XX 01-FEB-2002; 2002US-0353058P.

XX 04-MAR-2002; 2002US-0361293P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX (AMHP) WYETH.

XX Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;

XX
DR WPI; 2003-129214/12.
DR P-PSDB; ADB98703.

XX New nucleic acid comprising a mutation in LRP5 or LRP6, useful for
PT diagnosing a HBM-like phenotype in a subject and for preparing a
PT composition for modulating bone mass and/or lipid levels in a subject
PT suffering from e.g. osteoporosis.

XX Disclosure; SEQ ID NO 677; 629pp; English.

XX The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and
CC LRP6 mutants, which results in a HBM-like phenotype when expressed in a
CC cell. The HBM-like phenotype results in bone mass modulation and/or lipid
CC level modulation. The invention is useful for diagnosing a HBM-like
CC phenotype in a subject and for preparing a composition for modulating
CC bone mass and/or lipid levels in a subject suffering from e.g.
CC osteoporosis. The present sequence was used to illustrate the invention.

XX SQ Sequence 2328 AA;

Query Match 100.0%; Score 275; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
Db 5 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 53

RESULT 14

ADB82522
ID ADB82522 standard; protein; 2328 AA.

XX ADB82522;

XX 29-JAN-2004 (first entry)

XX Human protein sequence related to the invention #12.

XX LRP5; LRP6; HBM; Dkk activity; Osteopathic; Antiinflammatory;
KW Antiarthritic; bone mass disorders; osteoporosis; hypercalcaemia;
KW hyperostosis; osteogenesis; Wnt signaling.

XX Homo sapiens.

XX WO200292015-A2.

XX 21-NOV-2002.

XX 17-MAY-2002; 2002WO-US015982.

XX 17-MAY-2001; 2001US-0291311P.

XX 01-FEB-2002; 2002US-0353058P.

XX 04-MAR-2002; 2002US-0361293P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX (AMHP) WYETH.

XX Allen K, Anisowicz A, Bhat BM, Damagnez V, Robinson JA;

XX Yaworsky PJ;

XX WPI; 2003-129219/12.

XX Regulating LRP5, LRP6 or HBM activity in a subject, useful for modulating
PT lipid levels and/or bone mass, and for in treating bone mass disorders,
PT e.g. osteoporosis, comprises administering a composition which modulates
PT a Dkk activity.

XX Disclosure; SEQ ID NO 98; 173pp; English.

XX The present invention relates to regulating LRP5, LRP6 or HBM activity in
CC a subject comprising administering a composition which modulates a Dkk

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: September 30, 2005, 13:30:51 ; Search time 7.5627 Seconds
(without alignments)
623.404 Million cell updates/sec
Title: US-09-492-971b-15_COPY_4_52
Perfect score: 275
Sequence: 1 QAAQWVQPSPAVVSQSKPG.....INQWERTYLGNVLVCTCVG 49
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB ID	Description	
1	275	100.0	2386	1 FNHU	fibronectin precursor	
2	257	93.5	2477	2 S14428	fibronectin prec	
3	255	92.7	2265	1 FNBO	fibronectin - bovi	
4	194	70.5	2481	2 A43508	fibronectin - Afri	
5	110	40.0	103	2 A49173	fibronectin - mous	
6	79.5	28.9	1020	2 A29355	fibronectin - chic	
7	67.5	24.5	190	2 I51279	fibronectin - east	
8	57.5	20.9	941	1 A55195	chordin precursor	
9	57	20.7	586	1 SYMSAL	5-aminolevulinate	
10	57	20.7	587	1 SYHUA6	5-aminolevulinate	
11	56.5	20.5	272	2 S03412	hypothetical prote	
12	56.5	20.5	365	2 C97433	polyamine transpor	
13	56.5	20.5	338	2 AD2651	hypothetical prote	
14	55	20.0	338	2 T10544	inositol 1,3,4-tri	
15	55	20.0	645	2 T29818	hypothetical prote	
16	54.5	19.8	114	2 A34567	beta-microseminopr	
17	54.5	19.8	119	2 F60767	copia polyprotein	
18	54.5	19.8	169	2 A40522	plasma1n (EC 3.4.21	
19	54.5	19.8	220	2 T45056	hypothetical prote	
20	54.5	19.8	362	2 S24551	protein-tyrosine k	
21	54.5	19.8	482	2 T22754	hypothetical prote	
22	54.5	19.8	648	2 A48646	amine oxidase (cop	
23	54.5	19.8	787	2 T31232	copia polyprotein	
24	54.5	19.8	1170	1 TSHUP1	thrombospondin 1 p	
25	54.5	19.8	1409	1 OFFFCP	copia polyprotein	
26	54	19.6	149	2 C82800	hypothetical prote	
27	54	19.6	428	2 B84964	adenosylmethionine	
28	54	19.6	836	2 AD2565	hypothetical prote	
29	54	19.6	920	2 F71823	isoleucine-tRNA 11	

30	53.5	19.5	662	2 T46228	hypothetical prote
31	53	19.3	321	2 T24599	hypothetical prote
32	53	19.3	339	2 S47342	UDP-3-O- [3-hydroxy
33	53	19.3	341	2 D64102	UDP-3-O- [3-hydroxy
34	53	19.3	567	2 F64659	gamma-glutamyltran
35	53	19.3	567	2 H71856	gamma-glutamyltran
36	52.5	19.1	133	2 S75804	plastoquinol-plast
37	52.5	19.1	219	1 A39574	leukocyte antigen
38	52.5	19.1	554	2 S22495	pollen-specific pr
39	52.5	19.1	1170	2 A40558	thrombospondin 1 p
40	52	18.9	172	2 E82622	hypothetical prote
41	52	18.9	214	2 E75613	hypothetical prote
42	52	18.9	250	2 A42103	Distal-less (Dll)
43	52	18.9	472	2 AD2532	hypothetical prote
44	52	18.9	520	2 E90304	conserved hypochet
45	52	18.9	590	2 I39685	polyhydroxyalkanoi

ALIGNMENTS

RESULT 1

FNHU
fibronectin precursor [validated] - human
N:Alternate names: fibronectin splice form ED-A
C:Species: Homo sapiens (man)
C>Date: 27-NOV-1985 #sequence revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: A26460; A26284; A24854; A24476; A91008; A93529; A21011; A90495; A2
R:Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A:Reference number: A26460; MUID:87175578; PMID:3031656
A:Accession: A26460
A:Molecule type: DNA
A:Residues: 1-49 <DEA>
A:Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5533
R:Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A:Title: Evolution of the fibronectin gene.
A:Reference number: A26284; MUID:86111901; PMID:3003095
A:Accession: A26284
A:Molecule type: DNA
A:Residues: 1447-1540 <OLD>
A:Cross-references: GB:M12549; NID:g182688
A:Note: the authors translated the codon TTC for residue 1494 as Glu
R:Paolella, G.; Henschcliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-
A:Reference number: S00848; MUID:88233940; PMID:3375063
A:Accession: S03917
A:Molecule type: DNA
A:Residues: 1594-1767, 'V', 1769-1783 <PAO>
A:Cross-references: EMBL:X07718; NID:g31402
A:Note: the authors translated the codon AAC for residue 1631 as Asp
R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene
A:Reference number: A24854; MUID:87030929; PMID:3770201
A:Accession: A24854
A:Molecule type: DNA
A:Residues: 1992-2147 <VIB>
A:Cross-references: GB:X04530; NID:g31436
R:Gutman, A.; Yamada, K.M.; Kornblihtt, A.
FEBS Lett. 207, 145-148, 1986
A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
A:Reference number: A24476; MUID:87030890; PMID:3770189
A:Accession: A24476
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-14, 'O', 16-38 <GUT>
R:Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A:Title: Primary structure of human fibronectin: differential splicing may generate at

A:Reference number: A91008; MUID:85284965; PMID:2992939
A:Accession: A91008
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 32-1344,1346-2080;2112-2386 <KOR>
A:Cross-references: GB:X02761
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A:Reference number: A93529; MUID:84272258; PMID:6462919
A:Accession: A93529
A:Molecule type: mRNA
A:Residues: 973-2080;2112-2386 <K02>
A:Cross-references: GB:X00739
R:Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A:Reference number: A21011; MUID:83290929; PMID:6688418
A:Accession: A21011
A:Molecule type: mRNA
A:Residues: 1434-1537 <OL2>
A:Cross-references: GB:X00055; NID:g182680; PIDN:AAAS2459.1; PID:g182683
R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A:Reference number: A90495; MUID:85280409; PMID:2992573
A:Accession: A90495
A:Molecule type: mRNA
A:Residues: 1594-2386 <BER>
A:Cross-references: GB:M10905; NID:g182696; PIDN:AAAS2462.1; PID:g182697
R:Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A:Reference number: A22245; MUID:85231203; PMID:2989004
A:Accession: A22245
A:Molecule type: mRNA
A:Residues: 1948-2067 <UM5>
A:Cross-references: GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID:g182706
A:Accession: B22245
A:Molecule type: mRNA
A:Residues: 1975-1991;2017-2039 <UM2>
A:Cross-references: GB:M27590
R:Seikiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A:Title: Human liver fibronectin complementary DNAs: identification of two different me
A:Reference number: I52394; MUID:87026578; PMID:3021206
A:Accession: I65273
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SEK>
A:Cross-references: GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID:g182704
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
A:Reference number: A21165; MUID:83221567; PMID:6304699
A:Accession: A21165
A:Molecule type: mRNA
A:Residues: 2291-2386 <K03>
A:Cross-references: GB:X00799; NID:g182681; PIDN:AAAS2460.1; PID:g182684
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A:Title: Primary structure of human plasma fibronectin.
A:Reference number: A92398; MUID:84032463; PMID:663202
A:Accession: A92398
A:Molecule type: protein
A:Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
R:Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A:Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A:Reference number: S34791; MUID:93312001; PMID:8223285
A:Accession: S34791
A:Molecule type: protein
A:Residues: 291-300;551-560 <GAR2>

R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A:Reference number: A60904; MUID:87019725; PMID:3532418
A:Accession: A60904
A:Molecule type: protein
A:Residues: 293-301 <GRI>
R:Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A:Reference number: A23901; MUID:86008277; PMID:3900070
A:Accession: A23901
A:Molecule type: protein
A:Residues: 616-677,'O',679-703,'PT' <CAL>
R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A:Title: The cell attachment domain of fibronectin. Determination of the primary structu
A:Reference number: A92386; MUID:82265604; PMID:7050098
A:Accession: A92386
A:Molecule type: protein
A:Residues: 1441-1548 <PIE>
A:Note: residues 1524-1527 are responsible for the cell-binding activity
R:Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A:Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A:Reference number: A32517; MUID:87241275; PMID:3593230
A:Accession: A32517
A:Molecule type: protein
A:Residues: 1599-1630,'T',1722-2058 <GAR3>
R:Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A:Title: Human plasma fibronectin. Demonstration of structural differences between the A
A:Reference number: S14357; MUID:91190085; PMID:2012601
A:Accession: S14357
A:Molecule type: protein
A:Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A:Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A:Reference number: A23891; MUID:85261459; PMID:4019516
A:Accession: A23891
A:Molecule type: protein
A:Residues: 2071-2080;2112-2356 <GAR4>
R:Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
action, and transformation.
C:Genetics:
A:Gene: GDB:FN1
A:Cross-references: GDB:119135; OMIM:135600
A:Map position: 2q34-q34
A:Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-31/Domain: propeptide #status predicted <PRO>
F:32-2386/Product: fibronectin #status experimental <MAT>
F:52-272/Domain: fibrin and heparin binding <HPB>
F:52-87/Domain: fibronectin type I repeat homology <IF1>
F:97-135/Domain: fibronectin type I repeat homology <IF2>
F:141-179/Domain: fibronectin type I repeat homology <IF3>
F:186-225/Domain: fibronectin type I repeat homology <IF4>
F:231-270/Domain: fibronectin type I repeat homology <IF5>
F:308-608/Domain: collagen binding <CB>
F:308-342/Domain: fibronectin type I repeat homology <IF6>
F:360-401/Domain: fibronectin type II repeat homology <2F1>
F:420-461/Domain: fibronectin type II repeat homology <2F2>
F:470-508/Domain: fibronectin type I repeat homology <IF7>
F:518-555/Domain: fibronectin type I repeat homology <IF8>
F:561-599/Domain: fibronectin type I repeat homology <IF9>
F:609-692/Domain: fibronectin type III repeat homology <3FA>
F:616-706/Domain: heparin binding <HPB>
F:719-801/Domain: fibronectin type III repeat homology <3FB>
F:810-891/Domain: fibronectin type III repeat homology <3FC>

R;Khandjian, E.W.; Salomon, C.; Leonard, N.; Tremblay, S.; Turler, H.
Exp. Cell Res. 202, 464-470, 1992

A>Title: Fibronectin gene expression in proliferating, quiescent, and SV40-infected mouse fibroblasts
A:Reference number: A49173; MUID:93011702; PMID:1327855

A:Accession: A49173

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-103 <KHA>

A:Cross-references: UNIPROT:P11276; GB:S45680; NID:g256715; PIDN:AAB23491.1; PID:g2567156

A:Experimental source: kidney cells

A>Note: sequence extracted from NCBI backbon (NCBIN:115080, NCBIP:115081)

R;Blatti, S.P.; Foster, D.N.; Ranganathan, G.; Moses, H.L.; Getz, M.J.
Proc. Natl. Acad. Sci. U.S.A. 85, 1119-1123, 1988

A>Title: Induction of fibronectin gene transcription and mRNA is a primary response to growth factors

A:Reference number: A31371; MUID:98124987; PMID:3124113

A:Accession: A31371

A:Molecule type: mRNA

A:Residues: 1-65, 'N', 67-103 <BLA>

A:Cross-references: GB:M18194; GB:J03646; NID:gl93329; PIDN:AAA37636.1; PID:g387159

R;Ryseck, R.P.; Macdonald-Bravo, H.; Zerlal, M.; Bravo, R.
Exp. Cell Res. 180, 537-545, 1989

A>Title: Coordinate induction of fibronectin, fibronectin receptor, tropomyosin, and actin

A:Reference number: A60597; MUID:99121031; PMID:2521606

A:Accession: G60597

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-103 <RYS>

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology; alternative splicing; duplication; extracellular matrix; glycoprotein; heterodimer

F;11-46/Domain: fibronectin type I repeat homology <IFI2>
F;11-37,35-46/Disulfide bonds: #status predicted
F;84/Disulfide bonds: interchain (to 88) #status predicted
F;88/Disulfide bonds: interchain (to 84) #status predicted

Query Match 40.08; Score 110; DB 2; Length 103;
Best Local Similarity 51.79; Pred. No. 1.9e-06;
Matches 15; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 21 CYDNGKHVQIQNWERTYLGVLVCTCYG 49
||||| :||| :||| :||| :|||
Db 11 CYDDGRTHVGEQWQEYLGAICSTCFG 39

RESULT 6

A29355

fibronectin - chicken (fragment)

C:Species: Gallus gallus (chicken)

C>Date: 31-Dec-1998 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A29355

R;Norton, P.A.; Hynes, R.O.
Mol. Cell. Biol. 7, 4297-4307, 1987

A>Title: Alternative splicing of chicken fibronectin in embryos and in normal and transformed cells

A:Reference number: A29355; MUID:88142820; PMID:2830487

A:Accession: A29355

A:Molecule type: mRNA

A:Residues: 1-1020 <NOR>

A:Cross-references: UNIPROT:P11722

C:Genetics:

A:introns: 176/3

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology; alternative splicing; duplication; extracellular matrix; glycoprotein; heterodimer

F;1-86/Domain: fibronectin type III repeat homology <FN3H>

F;92-177/Domain: fibronectin type III repeat homology <FN3I>

F;180-262/Domain: fibronectin type III repeat homology <FN3J>

F;257-259/Region: cell attachment (R-G-D) motif

F;274-356/Domain: fibronectin type III repeat homology <FN3K>

F;364-446/Domain: fibronectin type III repeat homology <FN3L>

F;454-536/Domain: fibronectin type III repeat homology <FN3M>

F;546-628/Domain: fibronectin type III repeat homology <FN3N>

F;636-718/Domain: fibronectin type III repeat homology <FN3O>

F;837-917/Domain: fibronectin type III repeat homology <FN3P>

F;940-969/Domain: fibronectin type I repeat homology <FI10>

F;940-969,967-979,985-1012/Disulfide bonds: #status predicted

Query Match 28.9%; Score 79.5; DB 2; Length 1020;
Best Local Similarity 41.2%; Pred. No. 0.1;
Matches 14; Conservative 8; Mismatches 11; Indels 1; Gaps

Qy 17 SKFGCYDNGKHYYIQNWERT-YLGNVLVCTCYG 49
 | | | | | | | | | |
Db 981 SSKWCHDNGVNYKI GEKWDROENGOMIDTCYCLG 1014

RESULT 7
I51279
fibronectin - eastern newt (fragment)
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51279
R;Name, J.D.; Tassava, R.A.
Dev. Dyn. 202, 153-164, 1995
A;Title: Examination of fibronectin distribution and its sources in the reg
A;Reference number: I51279; UID:95252528; PMID:7734733
A;Accession: I51279
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-190 <NAC>

Query Match	24.5%	Score 67.5;	DB 2;	Length 190;
Best Local Similarity	38.5%;	Pred. No. 0.6;		
Matches	15; Conservative	8; Mismatches	15; Indels	1; Gaps 1;
Qy	12 VAVSQSKPGCYDNGKHVQINQOW-ERTYLGNNVLVCTCYG	49		
Db	96 VAVSOLRDCCIYDGITIVDVNHTTKRHEEGHMMNCTCYG	134		

```

RESULT 8
A55195
chordin precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A55195
R;Sasai, Y.; Lu, B.; Steinbeisser, H.; Geissert, D.; Gont, L.K.; De Roberti
Cell 79, 779-790, 1994
A;Title: Xenopus chordin: a novel dorsaling factor activated by organizer
A;Reference number: A55195; MUID:95094250; PMID:8001117
A;Accession: A55195
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-941 <SAS>

```

	Query Match	20.9%	Score 57.5;	DB 1;	Length 941;
	Best Local Similarity	31.9%	Pred. No. 49;		
	Matches	15;	Conservative	6;	Mismatches 21;
					Indels 5;
					Gaps 3;
Qy	6	VQPSQPVAVSQSPGCGYDNGKHQYQINQOWERTYLG	---	NVLVCT-CY	48
Db	28	IQEQEPLQSKTPAGCTFGGKPFYSLSDSWHPD-LGEPFGVMHCVLCY			73

RESULT 9
SYMSAL
5-aminolevulinate synthase (EC 2.3.1.37) precursor, erythroid-specific, mitochondrial

N;Alternate names: delta-aminolevulinate synthase, nonspecific
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004
C;Accession: A29040
R;Schoenhaut, D.S.; Curtis, P.J.
Gene 48, 55-63, 1986
A;Title: Nucleotide sequence of mouse 5-aminolevulinic acid synthase cDNA and expression
A;Reference number: A29040; MUID:87163516; PMID:3557128
A;Accession: A29040
A;Molecule type: mRNA
A;Residues: 1-586 <SCH>
A;Cross-references: UNIPROT:P08680; GB:M15268; NID:g191857; PIDN:AAA37207.1; PID:g387096
R;Ferreira, G.C.; Neame, P.J.; Dailey, H.A.
Protein Sci. 2, 1959-1965, 1993
A;Title: Heme biosynthesis in mammalian systems: evidence of a Schiff base linkage between
A;Reference number: A48183; MUID:94093402; PMID:8268805
A;Contents: annotation: cofactor binding site
C;Comment: The transit peptide is removed from the nuclear-encoded precursor during trans-
idoxal phosphate; it catalyzes the formation of 5-aminolevulinic acid from succinyl CoA
C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology
C;Keywords: acyltransferase; coenzyme A; mitochondrial matrix; mitochondrion; phosphop
F;1-8/Domain: transit peptide (mitochondrion) (fragment) #status predicted <TRM>
F;9-586/Product: 5-aminolevulinate synthase #status predicted <MAT>
F;193-333/Domain: glycine C-acetyltransferase homology <GCA>
F;390/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 20.7%; Score 57; DB 1; Length 586;
Best Local Similarity 43.8%; Pred. No. 36;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 27 HYQINQOQWERTYLGNV 42
|::: |||||::|
Db 562 HFELMSBWSRYSFGNM 577

RESULT 10
SYHVAE
5-aminolevulinate synthase (EC 2.3.1.37) precursor, erythroid-specific, mitochondrial -
N;Alternate names: 5-aminolevulinate synthase 2; erythroid delta-aminolevulinate synthase
C;Species: Homo sapiens (man)
C;Date: 21-Nov-1993 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C;Accession: S16347; PMID:1266919; PMID:2050125
EMBO J. 10, 1891-1902, 1991
A;Title: Human erythroid 5-aminolevulinate synthase: promoter analysis and identification
A;Reference number: S16347; MUID:91266919; PMID:2050125
A;Accession: S16347
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-587 <COX>
A;Cross-references: UNIPROT:P22557; EMBL:X60364; NID:g28587; PIDN:CAA42916.1; PID:g28587
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 2 in having 147-leu-
Nucleic Acids Res. 18, 7187-7188, 1990
A;Title: Two different genes encode delta-aminolevulinate synthase in humans: nucleotide
A;Reference number: S13682; MUID:91088347; PMID:2263504
A;Accession: S13683
A;Molecule type: mRNA
A;Residues: 6-181, 'F', 183-587 <BIS>
A;Cross-references: EMBL:X56352; NID:g28585; PIDN:CAA39795.1; PID:g28586
C;Genetics:
A;Gene: GDB:ALAS2; ASB: ALASE
A;Cross-references: GDB:119666; OMIM:301300
A;Map position: Xp11.21-Xp11.21
A;Note: defects in this gene may result in sideroblastic/hypochromic anemia
C;Complex: homodimer
C;Function:
A;Description: catalyzes the formation of 5-aminolevulinic acid from succinyl CoA and gl
A;Pathway: porphyrin biosynthesis
C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology
C;Keywords: acyltransferase; coenzyme A; homodimer; mitochondrial matrix; mitochondrion;
F;1-49/Domain: transit peptide (mitochondrion) #status predicted <TRP>
F;50-587/Product: 5-aminolevulinate synthase 2 #status predicted <MAT>

Db 138 G 138

RESULT 13

AD2651

hypothetical protein potF [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AD2651

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Cross-references: UNIPROT:Q8UHS0; GB:AE008688; PIDN:AAL41626.1; PID:g17738965; GSPDB:C58

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: potF

A:Map position: circular chromosome

C:Superfamily: Escherichia coli spermidine/putrescine-binding protein

Query Match 20.5%; Score 56.5; DB 2; Length 365;

Best Local Similarity 24.6%; Pred. No. 26;

Matches 15; Conservative 3; Mismatches 18; Indels 25; Gaps 1;

QY 5 MYQPSFVAVSQSKPC-----YDNGKHQYINQOWERTYL 39

Db 78 VVAPTSPFLAQIKAGVYQYKLSKLPNLKNWPDITERLAKYDPGNEAVNVMGTTGI 137

QY 40 G 40

Db 138 G 138

RESULT 14

T10544

inositol 1,3,4-trisphosphate 5/6-kinase homolog T12G13.10 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: T10544

R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft, submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533

A:Accession: T10544

A:Molecule type: DNA

A:Residues: 1-338 <BEV>

A:Cross-references: UNIPROT:Q8UG3; EMBL:AL080252; GSPDB:GN00062; ATSP:T12G13.10

A:Experimental source: cultivar Columbia; BAC clone T12G13

C:Genetics:

A:Gene: ATSP:T12G13.10

A:Map position: 4

A:Introns: 65/2; 92/3; 103/3; 144/1; 175/1; 215/1; 274/1; 292/3

Query Match 20.0%; Score 55; DB 2; Length 338;

Best Local Similarity 41.4%; Pred. No. 37;

Matches 12; Conservative 5; Mismatches 8; Indels 4; Gaps 1;

QY 12 VAVSQSKP---GCDNGKHQYINQOWER 36

Db 72 VAIQNKPLSEQGFDFVLHKQIGKEWR 100

RESULT 15

T29818

hypothetical protein F25E2.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T29818

R:Minx, P.; Le, T.

A:Description: The sequence of C. elegans cosmid F25E2.

A:Reference number: Z20691

A:Accession: T29818

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-645 <MIN>

A:Cross-references: UNIPROT:Q19780; EMBL:U50197; PIDN:AAA91255.1; CESP:F25E2.2

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F25E2.2

A:Introns: 54/1; 94/1; 154/1; 179/3; 209/1; 249/1; 321/1; 422/1; 551/2; 623/1

Query Match 20.0%; Score 55; DB 2; Length 645;

Best Local Similarity 38.5%; Pred. No. 70;

Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 20 GCDNGKHQYINQOWERTYLGNVLC 45

Db 251 GCRHNGKVYRHNENWSEDCATVYAC 276

Search completed: September 30, 2005, 13:31:51

Job time : 9.5627 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 13:30:52 ; Search time 35.4502 Seconds
(without alignments)
707.807 Million cell updates/sec

Title: US-09-492-971b-15_COPY_4_52

Perfect score: 275
Sequence: 1 QAQMVPQSPVAVSQSPG.....INQWERTYLGNIIVCTCYG 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	271	98.5	749	2	Q7Z391
2	271	98.5	1103	2	Q6MZF4
3	271	98.5	2296	2	Q6NOA6
4	271	98.5	2357	2	Q68DT4
5	271	98.5	2386	1	FINC_HUMAN
6	271	98.5	2444	2	Q6N025
7	271	98.5	2477	2	Q6MZWU
8	267	97.1	2267	2	Q68DP9
9	263	95.6	296	2	Q8C6J7
10	263	95.6	2477	1	FINC_MOUSE
11	262	95.3	2240	2	Q68DF8
12	257	93.5	2477	1	FINC_RAT
13	255	92.7	2265	1	FINC_BOVIN
14	194	70.5	2481	1	FINC_XENLA
15	194	70.5	2481	2	Q6GOA5
16	191	69.5	215	2	Q6DD34
17	120	43.6	358	2	Q7T1S2
18	117	42.5	2478	2	O93406
19	112	40.7	922	2	O93405
20	110.5	40.2	2408	2	Q6JAN2
21	110	40.0	100	1	FINC_RABIT
22	110	40.0	111	2	Q862C9
23	110	40.0	163	2	Q9H382
24	110	40.0	211	2	Q7L553
25	110	40.0	216	2	Q99KD0
26	110	40.0	268	2	Q6PJE5
27	110	40.0	412	2	Q7L1U4
28	110	40.0	810	2	Q8R3F3
29	110	40.0	1034	2	Q6N084
30	110	40.0	1255	2	Q6MZWU
31	110	40.0	2146	2	Q68CX6

32	110	40.0	2193	2	Q6MZW7	Q6mzm7 homo sapien
33	104	37.8	351	2	O02816	O02816 oryctolagus
34	103.5	37.6	1328	1	FINC_PLEWA	Q91289 pleurodeles
35	87.5	31.8	347	2	Q95KV5	Q95kv5 bos taurus
36	82.5	30.0	195	2	Q7T2W7	Q7t2w7 brachydanio
37	80.5	29.3	62	2	Q28693	Q28693 oryctolagus
38	80.5	29.3	320	2	Q95KV4	Q95kv4 bos taurus
39	80.5	29.3	522	1	FINC_CANFA	Q28275 canis faml
40	80.5	29.3	522	1	FINC_HORSE	Q28377 equus cabal
41	79.5	28.9	1256	1	FINC_CHICK	P11722 gallus gall
42	67.5	24.5	190	1	FINC_NOTVI	Q91400 notophthalm
43	67	24.4	552	2	O7QNK5	O7qnk5 anopheles g
44	65.5	23.8	141	2	O90XQ2	O90xq2 ambystoma m
45	64.5	23.5	2768	2	Q9VC00	Q9vc00 drosophila

ALIGNMENTS

RESULT 1

Q7Z391	PRELIMINARY;	PRT;	749 AA.
AC Q7Z391;			
DT 01-OCT-2003 (Tremblrel. 25, Created)			
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)			
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE Hypothetical protein DKFZp686B18150.			
GN Name=DKFZp686B18150;			
OS Homo sapiens (Human);			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Human colon endothel primary cell culture;			
RA Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,			
RA Fobo G., Han M., Wiemann S.;			
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.			
DR EMBL; BX538045; CAD97984.1; -			
DR HSP; O96KP7; 1FBR.			
DR GO; GO:0005576; C:extracellular; IEA.			
DR InterPro; IPR006209; EGF like.			
DR InterPro; IPR000083; Fibrnctnl.			
DR Pfam; PF00039; fnl; 9.			
DR Pfam; PF00040; fn2; 2.			
DR ProDom; PD000395; FN_Type_II; 2.			
DR SMART; SM00058; FN1; 9.			
DR SMART; SM00059; FN2; 2.			
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.			
DR PROSITE; PS01253; FIBRONECTIN_1; 9.			
DR PROSITE; PS00023; FIBRONECTIN_2; 2.			
KW Hypothetical protein.			
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;			

Query Match 98.5%; Score 271; DB 2; Length 749;
Best Local Similarity 98.0%; Pred. No. 3.7e-25;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	QAQMVPQSPVAVSQSPGCGYDNGKHQYINQWERTYLGNIIVCTCYG 49
DB	124	QAQMVPQSPVAVSQSPGCGYDNGKHQYINQWERTYLGNIIVCTCYG 172

RESULT 2

Q6MZF4	PRELIMINARY;	PRT;	1103 AA.
ID Q6MZF4;			
AC Q6MZF4;			
DT 05-JUL-2004 (Tremblrel. 27, Created)			
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)			
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)			
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219) (Fragment).			

DR	InterPro; IPR000562; FN_Type_II.
DR	Pfam; PF00039; fn1; 12.
DR	Pfam; PF00040; fn2; 2.
DR	Pfam; PF00041; fn3; 15.
DR	PRINTS; PR00012; FNTYPEI.
DR	PRINTS; PR00013; FNTYPEII.
DR	ProDom; PD000995; FN_Type_II; 2.
DR	SMART; SM00058; FN1; 12.
DR	SMART; SM00059; FN2; 2.
DR	SMART; SM00060; FN3; 15.
DR	PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR	PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR	PROSITE; PS01253; FIBRONECTIN_1; 12.
DR	PROSITE; PS00023; FIBRONECTIN_2; 2.
DR	PROSITE; PS00023; FIBRONECTIN_2; 2.
DR	Hypothetical protein.
KW	SEQUENCE 2296 AA; 252761 MW; 9AB2D723CC0CED70 CRC64;
SQ	
Query Match	98.5%; Score 271; DB 2; Length 2296;
Best Local Similarity	98.0%; Pred. No. 1.2e-24;
Matches	48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 QAAQMVPQSPVAVSQSGPCVDNGKHQIQINQWERTYLGNVLVCTCYG 49
DB	32 QAAQMVPQSPVAVSQSGPCVDNGKHQIQINQWERTYLGNALVCTCYG 80
RESULT 4	
Q68DT4	
ID	Q68DT4 PRELIMINARY; PRT; 2357 AA.
AC	Q68DT4;
DT	25-OCT-2004 (TrEMBLrel. 28, Created)
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Hypochemical protein DKFPz686Ff0164.
OS	Name=DKFPz686Ff0164;
OC	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NB	NCBI_TaxID=9606;
RN	[1]_TaxID=9606;
RP	SEQUENCE FROM N.A.
RC	TISSUE=uterus endothel;
RG	The German cDNA Consortium;
RA	Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RL	Fobo G., Han M., Wiemann S.;
RU	Submitted [AUG-2004] to the EMBL/GenBank/DBJ databases.
DR	EMBL; CX749281; CAH18136.1; -
DR	InterPro; IPR002086; Aldehyd dehydrog.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; Fibrinctnl.
DR	InterPro; IPR003962; FnIII subd.
DR	InterPro; IPR003961; FN III.
DR	InterPro; IPR008957; FN_III-like.
DR	InterPro; IPR000582; FN_Type_II.
DR	Pfam; PF00039; fn1; 12.
DR	Pfam; PF00040; fn2; 2.
DR	Pfam; PF00041; fn3; 15.
DR	PRINTS; PR00012; FNTYPEI.
DR	PRINTS; PR00013; FNTYPEII.
DR	PRINTS; PR00014; FNTYPEIII.
DR	ProDom; PD000995; FN_Type_II; 2.
DR	SMART; SM00058; FN1; 12.
DR	SMART; SM00059; FN2; 2.
DR	SMART; SM00060; FN3; 15.
DR	PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR	PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR	PROSITE; PS01253; FIBRONECTIN_1; 12.
DR	PROSITE; PS00023; FIBRONECTIN_2; 2.
KW	Hypothetical protein.
SQ	SEQUENCE 2357 AA; 259090 MW; BEAE3990E27E532A CRC64;
Query Match	98.5%; Score 271; DB 2; Length 2357;

RT HepG2.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
RN [19]
RP O-GLYCOSYLATION OF THR-2064.
RX MEDLINE=91190085; PubMed=2012601;
RA Tresselt T., McCarthy J.B., Calaycay J., Lee T.D., Legesse K.,
RA Shively J.E., Pande H.;
RT "Human plasma fibronectin. Demonstration of structural differences
RT between the A- and B-chains in the III CS region.";
RL Biochem. J. 274:731-738(1991).
RN [20]
RP FBLN1-BINDING SITE.
RX MEDLINE=93015979; PubMed=1400330;
RA Balbona K., Tran H., Godyna S., Ingham K.C., Strickland D.K.,
RA Agraves W.S.;
RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding
RT region of fibronectin.";
RL J. Biol. Chem. 267:20120-20125(1992).
RN [21]
RP CHARACTERIZATION OF FIBRIN-BINDING SITE 1.
RX MEDLINE=93081153; PubMed=7989369;
RA Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;
RT "Further characterization of the NH2-terminal fibrin-binding site on
RT fibronectin.";
RL J. Biol. Chem. 269:31938-31945(1994).
RN [22]
RP INTERACTION WITH LGALS3BP.
RX PubMed=9501082; DOI=10.1093/emboj/17.6.1606;
RA Sasaki T., Brakebusch C., Engel J., Timpl R.;
RT "Mac-2 binding protein is a cell-adhesive protein of the extracellular
RT matrix which self-assembles into ring-like structures and binds beta1
RT integrins, collagens and fibronectin.";
RL EMBO J. 17:1606-1613(1998).
RN [23]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE=92162710; PubMed=1311202;
RA Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,
RA Campbell I.D.;
RT "1H NMR assignment and secondary structure of the cell adhesion type
RT III module of fibronectin.";
RL Biochemistry 31:2068-2073(1992).
RN [24]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE=93046665; PubMed=1423622; DOI=10.1016/0092-8674(92)90600-H;
RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;
RT "The three-dimensional structure of the tenth type III module of
RT fibronectin: an insight into RGD-mediated interactions.";
RL Cell 71:671-678(1992).
RN [25]
RP STRUCTURE BY NMR OF 182-275.
RX MEDLINE=94141923; PubMed=8308892;
RA Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,
RA Campbell I.D.;
RT "Solution structure of a pair of fibronectin type 1 modules with
RT fibrin binding activity.";
RL J. Mol. Biol. 235:1302-1311(1994).
RN [26]
RP STRUCTURE BY NMR OF 32-92.
RX MEDLINE=96069779; PubMed=7583666;
RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;
RT "High-resolution structural studies of the factor XIIIa crosslinking
RT site and the first type 1 module of fibronectin.";
RL Nat. Struct. Biol. 2:946-950(1995).
RN [27]
Query Match 98.5%; Score 271; DB 1; Length 2386;
Best Local Similarity 98.0%; Pred. No. 1.2e-24;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 QAOQMVQPSVAVSQSGKPGCYDNGKHQYINQWERTYLGNAVLCCTCYG 49
DB 32 QAOQMVQPSVAVSQSGKPGCYDNGKHQYINQWERTYLGNAVLCCTCYG 80

RESULT 6
Q6N025 PRELIMINARY; PRT; 2444 AA.
ID Q6N025
AC Q6N025;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M2451 (Fragment).
GN Name=DKFZp686M2451;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RG The German Human cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BX640731; CA645847.1; -;
DR GO: GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR Pfam; PF00039; fnl; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 16.
DR PRINTS; PR00013; ENTPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 16.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS50853; FN3; 16.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2444 AA; 268676 MW; 71C5E8C56A84C7BC CRC64;
Query Match 98.5%; Score 271; DB 2; Length 2444;
Best Local Similarity 98.0%; Pred. No. 1.3e-24;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 QAOQMVQPSVAVSQSGKPGCYDNGKHQYINQWERTYLGNAVLCCTCYG 49
DB 121 QAOQMVQPSVAVSQSGKPGCYDNGKHQYINQWERTYLGNAVLCCTCYG 169
RESULT 7
Q6MZUS PRELIMINARY; PRT; 2477 AA.
ID Q6MZUS
AC Q6MZUS;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686O1166.
GN Name=DKFZp686O1166;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RG The German Human cDNA Consortium;
RA Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,

RA Fobo G., Han M., Wiemann S.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX640875; CAE45932.1; -;
 DR GO; GO:0005576; C:extracellular; IEA.
 DR InterPro; IPR002086; Aldehyd dehydrog.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrnctnl.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR000562; FN_Type_II.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 17.
 DR PRINTS; PR00013; FNTYPEII.
 DR ProDom; PD000995; FN_Type_II; 2.
 DR SMART; SM00058; FN1; 12.
 DR SMART; SM00059; FN2; 2.
 DR SMART; SM00060; FN3; 17.
 DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
 DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.
 DR PROSITE; PS00853; FN3; 17.
 KW Hypothetical protein.
 SQ SEQUENCE 2477 AA; 272335 MW; D358D85C6B18207C CRC64;

Query Match 98.5%; Score 271; DB 2; Length 2477;
 Best Local Similarity 98.0%; Pred. No. 1.3e-24;
 Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OQAMVQVQSPVAVSQSGKGYDNGKHQYQINQOWERTYLGVLVCTCYG 49
 DB 32 OQAMVQVQSPVAVSQSGKGYDNGKHQYQINQOWERTYLGVLVCTCYG 80

RESULT 8

Q68DP9
 ID Q68DP9 PRELIMINARY; PRT; 2267 AA.
 AC Q68DP9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein DKFP686K08164.
 GN Name=DKFP686K08164;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus endothel;
 RG The German cDNA Consortium;
 RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
 RA Fobo G., Han M., Wiemann S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR749316; CAH18171.1; -;
 DR InterPro; IPR002086; Aldehyd dehydrog.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrnctnl.
 DR InterPro; IPR003962; FNIII subd.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR000562; FN_Type_II.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 16.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00014; FNTYPEIII.
 DR ProDom; PD000995; FN_Type_II; 2.
 DR SMART; SM00058; FN1; 12.
 DR SMART; SM00059; FN2; 2.
 DR SMART; SM00060; FN3; 16.

DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
 DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 2267 AA; 249358 MW; C4D124A038C323DF CRC64;

Query Match 97.1%; Score 267; DB 2; Length 2267;
 Best Local Similarity 95.9%; Pred. No. 3.7e-24;
 Matches 47; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 OQAMVQVQSPVAVSQSGKGYDNGKHQYQINQOWERTYLGVLVCTCYG 49
 DB 32 OQAMVQVQSPVAVSQSGKGYDNGKHQYQINQOWERTYLGVLVCTCYG 80

RESULT 9

Q8C6J7
 ID Q8C6J7 PRELIMINARY; PRT; 296 AA.
 AC Q8C6J7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330027I09 product:fibronectin 1, full insert sequence.
 DE Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayata N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kusunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format

CC synthesis.
 CC -!- PTM: Sulfated (By similarity).
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; BC051082; AAH51082.1; -
 CC EMBL; 222729; CAA80422.1; -
 CC EMBL; X82402; CAA57796.1; -
 CC EMBL; X91167; CAA63654.1; -
 CC EMBL; M18194; AAA37636.1; -
 CC EMBL; S45680; AAB23491.1; -
 CC PIR; A49173; A49173.
 CC PIR; I48349; I48349.
 CC PDB; 1MFN; NMR; @=1446-1630.
 CC PDB; 2MFN; NMR; @=1446-1630.
 CC MGD; MGI:95566; Fnl.
 CC GO; GO:0007155; P:cell adhesion; IDA.
 CC GO; GO:0007044; P:cell-substrate junction assembly; IDA.
 CC GO; GO:0042060; P:wound healing; IMP.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000083; Fibrinctnl.
 CC InterPro; IPR003982; FnlII subd.
 CC InterPro; IPR003961; FN_III.
 CC InterPro; IPR008957; FN_III-like.
 CC InterPro; IPR000562; FN_Type_II.
 CC Pfam; PF00039; fn1; 12.
 CC Pfam; PF00040; fn2; 2.
 CC Pfam; PF00041; fn3; 17.
 CC PRINTS; PR00014; FNTYPEIII.
 CC ProDom; PD000995; FN_Type_II; 2.
 CC PROSITE; PS00022; EGF 1; 2.
 CC PROSITE; PS01253; FIBRONECTIN 1; 12.
 CC PROSITE; PS00023; FIBRONECTIN_2; 2.
 CC PROSITE; PS00853; FN3; 17.
 CC 3D-structure; Acute phase; Alternative splicing; Cell adhesion;
 CC Glycoprotein; Heparin-binding; Phosphorylation; Plasma; Repeat;
 CC Signal; Sulfation.
 FT CHAIN 1 32 By similarity.
 FT SIGNAL 33 2477 Fibronectin.
 FT DOMAIN 53 273 Fibrin- and heparin-binding 1.
 FT DNA_BIND 308 608 Collagen-binding.
 FT DOMAIN 906 1171 Cell-attachment.
 FT DOMAIN 1357 1630 Heparin-binding 2.
 FT DOMAIN 1811 2081 Fibrin-binding 2.
 FT DOMAIN 2296 2427 Fibronectin type-I 1.
 FT DOMAIN 51 96 Fibronectin type-I 2.
 FT DOMAIN 96 140 Fibronectin type-I 3.
 FT DOMAIN 140 185 Fibronectin type-I 4.
 FT DOMAIN 185 230 Fibronectin type-I 5.
 FT DOMAIN 230 272 Fibronectin type-I 6.
 FT DOMAIN 306 343 Fibronectin type-II 1.
 FT DOMAIN 345 404 Fibronectin type-II 2.
 FT DOMAIN 405 469 Fibronectin type-II 3.
 FT DOMAIN 468 516 Fibronectin type-I 7.
 FT DOMAIN 516 559 Fibronectin type-I 8.
 FT DOMAIN 559 602 Fibronectin type-I 9.
 FT DOMAIN 607 699 Fibronectin type-III 1.
 FT DOMAIN 719 808 Fibronectin type-III 2.
 FT DOMAIN 810 897 Fibronectin type-III 3.
 FT DOMAIN 905 994 Fibronectin type-III 4.
 FT DOMAIN 995 1083 Fibronectin type-III 5.
 FT DOMAIN 1091 1171 Fibronectin type-III 6.
 FT DOMAIN 1172 1264 Fibronectin type-III 7.
 FT DOMAIN 1265 1355 Fibronectin type-III 8 (extra domain 1).

FT DOMAIN 1356 1446 Fibronectin type-III 9.
 FT DOMAIN 1447 1536 Fibronectin type-III 10.
 FT DOMAIN 1537 1626 Fibronectin type-III 11.
 FT DOMAIN 1631 1720 Fibronectin type-III 12.
 FT DOMAIN 1721 1810 Fibronectin type-III 13 (extra domain 2).
 FT DOMAIN 1813 1900 Fibronectin type-III 14.
 FT DOMAIN 1903 1991 Fibronectin type-III 15.
 FT DOMAIN 1992 2081 Fibronectin type-III 16.
 FT DOMAIN 2082 2201 Connecting strand 3 (CS-3) (V region).
 FT DOMAIN 2202 2280 Fibronectin type-III 17.
 FT DOMAIN 2281 2338 Fibronectin type-I 10.
 FT DOMAIN 2339 2381 Fibronectin type-I 11.
 FT DOMAIN 2383 2426 Fibronectin type-I 12.
 FT SITE 1614 1616 Cell attachment site.
 FT SITE 2181 2183 Cell attachment site.
 FT DISULFID 53 79 By similarity.
 FT DISULFID 77 88 By similarity.
 FT DISULFID 98 126 By similarity.
 FT DISULFID 124 136 By similarity.
 FT DISULFID 142 170 By similarity.
 FT DISULFID 168 180 By similarity.
 Query Match 95.6%; Score 263; DB 1; Length 2477;
 Best Local Similarity 93.9%; Pred. No. 1.3e-23;
 Matches 46; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QAQQVQPSVAVSQSPGCGYDNGKHQIQINQWERTYLGVLVCTCYG 49
 DB 33 QAQQVQPSVAVSQSPGCGYDNGKHQIQINQWERTYLGVLVCTCYG 81
 RESULT 11
 Q68DP8 PRELIMINARY; PRT; 2240 AA.
 AC Q68DP8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein DKFZp686H0342.
 GN Name=DKFZp686H0342;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Amygdala;
 RG The German CDNA Consortium;
 RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schapp A.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR749317; CAH18172.1; -
 DR InterPro; IPR002086; Aldehyd dehydrog.
 DR InterPro; IPR006209; EGF like
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR003961; FN_III subd.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR000562; FN_Type_II.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 15.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00014; FNTYPEIII.
 DR ProDom; PD000995; FN_Type_II; 2.
 DR SMART; SM00056; FN1; 12.
 DR SMART; SM00059; FN2; 2.
 DR SMART; SM00060; FN3; 15.
 DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
 DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
 DR PROSITE; PS01253; FIBRONECTIN 1; 10.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.

```

KW Hypothetical protein.
SQ SEQUENCE 2240 AA; 246667 MW; 8FCDAF406F330621 CRC64;

Query Match 95.3%; Score 262; DB 2; Length 2240;
Best Local Similarity 95.9%; Pred. No. 1.5e-23;
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QAAQMVQSPVAVSQSPGCGYDNGKHQYQINQWERTYLGNVLVCTCYG 49
   |||||
DQ 32 QAAQMVQSPVAVSQSPGCGYDNGKHQYQINQWERTYLGNVLVCTCYG 80

RESULT 12
ID_FINC_RAT STANDARD; PRT; 2477 AA.
AC P04937; Q6LXD9;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Fibronectin precursor (FN).
GN Name=Fnl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=88054951; PubMed=2445560;
RA Schwarzbauer J.E., Patel R.S., Fonda D., Hynes R.O.;
RT "Multiple sites of alternative splicing of the rat fibronectin gene
   transcript."
RL EMBO J. 6:2573-2580(1987).
RN [2]
RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=88054950; PubMed=3119323;
RA Patel R.S., Odermatt E., Schwarzbauer J.E., Hynes R.O.;
RT "Organization of the fibronectin gene provides evidence for exon
   shuffling during evolution."
RL EMBO J. 6:2565-2572(1987).
RN [3]
RP SEQUENCE OF 1586-2477 FROM N.A.
RX MEDLINE=84082067; PubMed=6317187; DOI=10.1016/0092-8674(83)90175-7;
RA Schwarzbauer J.E., Tamkun J.W., Lemischka I.R., Hynes R.O.;
RT "Three different fibronectin mRNAs arise by alternative splicing
   within the coding region."
RL Cell 35:421-431(1983).
RN [4]
RP SEQUENCE OF 1722-1810 FROM N.A., AND ALTERNATIVE SPLICING.
RX PubMed=3863113;
RA Odermatt E., Tamkun J.W., Hynes R.O.;
RT "Repeating modular structure of the fibronectin gene: relationship to
   protein structure and subunit variation."
RL Proc. Natl. Acad. Sci. U.S.A. 82:6571-6575(1985).
RN [5]
RP SEQUENCE OF 2052-2237 FROM N.A., AND ALTERNATIVE SPLICING.
RX PubMed=6089177;
RA Tamkun J.W., Schwarzbauer J.E., Hynes R.O.;
RT "A single rat fibronectin gene generates three different mRNAs by
   alternative splicing of a complex exon."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5140-5144(1984).
RN [6]
RP SEQUENCE OF 1183-1192; 1385-1399 AND 2287-2300, AND INTERACTION WITH
   AMBP.
RX PubMed=7519849;
RA Falkenberg C., Enghild J.J., Thøgersen I.B., Salvesen G.,
RA Akerstrom B.;
RT "Isolation and characterization of fibronectin-alpha 1-microglobulin
   complex in rat plasma."
RL Biochem. J. 301:745-751(1994).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
   including collagen, fibrin, heparin, DNA, and actin. Fibronectins

```

```

CC are involved in cell adhesion, cell motility, opsonization, wound
CC healing, and maintenance of cell shape.
CC -!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
CC variants, connected by 2 disulfide bonds near the carboxyl ends;
CC to a lesser extent homodimers. Interacts with FBLN1, AMBP and
CC LGALS3BP (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Comment=Each of the "extra domain" and the connecting strand 3
CC are present in some forms of fibronectin and absent in others;
CC Name=1;
CC IsoId=P04937-1; Sequence=Displayed;
CC Name=2; Synonyms=FNIII-13-less;
CC IsoId=P04937-2; Sequence=VSP_003258;
CC Name=3; Synonyms=Lambda-RLF4-5;
CC IsoId=P04937-3; Sequence=VSP_003259;
CC Name=4; Synonyms=Lambda-RLF6;
CC IsoId=P04937-4; Sequence=VSP_003260;
CC -!- TISSUE SPECIFICITY: Plasma FN (dimeric form) is secreted
CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
CC forms), made by fibroblasts, epithelial and other cell types, is
CC deposited as fibrils in the extracellular matrix.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15906; CAA34020.1; -
CC EMBL; L29191; AAA41166.1; -
CC EMBL; L00191; AAA41166.1; JOINED.
CC EMBL; L29191; AAA41167.1; -
CC EMBL; L00191; AAA41167.1; JOINED.
CC EMBL; L29191; AAA41168.1; -
CC EMBL; L00191; AAA41168.1; JOINED.
CC EMBL; M11750; AAA41170.1; -
CC EMBL; X05831; CAA29278.1; -
CC EMBL; X05832; CAA29279.1; -
CC EMBL; X05833; CAA29280.1; -
CC EMBL; X05834; CAA29281.1; -
CC PIR; S14428; S14428.
CC HSP; P08253; 1K50.
CC RGD; 2624; Fnl.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; FibrnctnI.
CC InterPro; IPR003962; FNIII_subd.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR008957; FN_III-like.
CC InterPro; IPR000562; FN_Type_II.
CC Pfam; PF00039; fni; 12.
CC Pfam; PF00040; fn2; 2.
CC Pfam; PF00041; fn3; 17.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00014; FNTYPEIII.
CC PRODOM; PD000995; FN_Type_II; 2.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01253; FIBRONECTIN_1; 12.
CC PROSITE; PS00023; FIBRONECTIN_2; 2.
CC PROSITE; PS00853; FN3; 17.
CC Acute phase; Alternative splicing; Cell adhesion;
CC Direct protein sequencing; Glycoprotein; Heparin-binding;
CC Phosphorylation; Plasma; Repeat; Signal; Sulfation.
FT SIGNAL 1
FT CHAIN 33 2477
FT DOMAIN 53 273
FT Fibrin- and heparin-binding 1.

```


FT	DOMAIN	308	608		Collagen-binding.	
FT	DNA_BIND	906	1171		Cell-attachment.	
FT	DOMAIN	1357	1630		Heparin-binding 2.	
FT	DOMAIN	1811	2081		Fibrin-binding 2.	
FT	DOMAIN	2296	2427		Fibronectin type-I 1.	
FT	DOMAIN	51	91		Fibronectin type-I 2.	
FT	DOMAIN	96	139		Fibronectin type-I 3.	
FT	DOMAIN	140	183		Fibronectin type-I 4.	
FT	DOMAIN	185	229		Fibronectin type-I 5.	
FT	DOMAIN	230	274		Fibronectin type-II 1.	
FT	DOMAIN	306	345		Fibronectin type-II 2.	
FT	DOMAIN	345	404		Fibronectin type-I 7.	
FT	DOMAIN	405	469		Fibronectin type-I 8.	
FT	DOMAIN	468	511		Fibronectin type-I 9.	
FT	DOMAIN	516	558		Fibronectin type-III 1.	
FT	DOMAIN	559	602		Fibronectin type-III 2.	
FT	DOMAIN	607	699		Fibronectin type-III 3.	
FT	DOMAIN	719	808		Fibronectin type-III 4.	
FT	DOMAIN	810	897		Fibronectin type-III 5.	
FT	DOMAIN	907	994		Fibronectin type-III 6.	
FT	DOMAIN	995	1083		Fibronectin type-III 7.	
FT	DOMAIN	1091	1171		Fibronectin type-III 8.	
FT	DOMAIN	1172	1264		Fibronectin type-III 9.	
FT	DOMAIN	1265	1355		Fibronectin type-III 10.	
FT	DOMAIN	1356	1446		Fibronectin type-III 11.	
FT	DOMAIN	1447	1536		Fibronectin type-III 12.	
FT	DOMAIN	1537	1626		Fibronectin type-III 13.	
FT	DOMAIN	1631	1720		Fibronectin type-III 14.	
FT	DOMAIN	1721	1810		Fibronectin type-III 15.	
FT	DOMAIN	1813	1900		Fibronectin type-III 16.	
FT	DOMAIN	1903	1991		Connecting strand 3 (CS-3) (V region).	
FT	DOMAIN	1992	2081		Fibronectin type-I 10.	
FT	DOMAIN	2082	2280		Fibronectin type-I 11.	
FT	DOMAIN	2190	2280		Fibronectin type-I 12.	
FT	DOMAIN	2294	2338		Fibronectin type-I 13.	
FT	DOMAIN	2339	2381		Fibronectin type-I 14.	
FT	DOMAIN	2383	2426		Fibronectin type-I 15.	
FT	SITE	1614	1616		Cell attachment site.	
FT	SITE	2181	2183		By similarity.	
FT	DISULFID	53	79		By similarity.	
FT	DISULFID	77	88		By similarity.	
FT	DISULFID	98	126		By similarity.	
FT	DISULFID	124	136		By similarity.	
FT	DISULFID	142	170		By similarity.	
FT	DISULFID	168	180		By similarity.	
FT	DISULFID	187	216		By similarity.	
FT	DISULFID	214	226		By similarity.	
FT	DISULFID	232	261		By similarity.	
FT	DISULFID	259	271		By similarity.	
FT	DISULFID	308	335		By similarity.	
FT	DISULFID	333	342		By similarity.	
FT	DISULFID	360	386		By similarity.	
FT	DISULFID	374	401		By similarity.	
FT	DISULFID	420	446		By similarity.	
FT	DISULFID	434	461		By similarity.	
FT	DISULFID	470	498		By similarity.	
FT	DISULFID	496	508		By similarity.	
FT	DISULFID	518	545		By similarity.	
FT	DISULFID	543	555		By similarity.	
FT	DISULFID	561	589		By similarity.	
FT	DISULFID	587	599		By similarity.	
FT	DISULFID	2296	2325		By similarity.	
FT	DISULFID	2323	2335		By similarity.	
FT	DISULFID	2341	2368		By similarity.	
FT	DISULFID	2366	2378		By similarity.	
FT	DISULFID	2385	2409		By similarity.	
FT	DISULFID	2407	2423		By similarity.	
FT	DISULFID	2458	2458		Interchain (with C-2462).	
Query Match					93.5%;	Score 257; DB 1; Length 2477;
Best Local Similarity					91.8%;	Pred. No. 7.2e-23;
Matches 45; Conservative 2;					Mismatches 2;	Indels 0; Gaps 0;

QY	1	QAQQWVQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG	49
DB	33	QAQQIVQPPSPVAVSQSKPGCFDNGKHQYQINQWERTYLGNVLVCTCYG	81
RESULT 13			
ID	FINC_BOVIN	STANDARD;	PRT; 2265 AA.
AC	P07589;		
DT	01-APR-1988	(Rel. 07, Created)	
DT	01-FEB-1994	(Rel. 28, Last sequence update)	
DT	05-JUL-2004	(Rel. 44, Last annotation update)	
DE	Fibronectin (FN).		
GN	Name=FN1;		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=87054047; PubMed=3780752;		
RA	Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;		
RT	"Complete primary structure of bovine plasma fibronectin.";		
RL	Eur. J. Biochem. 161:441-453(1986).		
RN	[2]		
RP	PARTIAL SEQUENCE.		
RX	MEDLINE=83117805; PubMed=6218503;		
RA	Petersen T.E., Thorgersen H.C., Skorstengaard K., Vibe-Pedersen K.,		
RA	Sahl P., Sottrup-Jensen L., Magnusson S.;		
RT	"Partial primary structure of bovine plasma fibronectin: three types		
RT	of internal homology.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).		
RN	[3]		
RP	SEQUENCE OF 2170-2265 FROM N.A.		
RX	MEDLINE=83221567; PubMed=6304699;		
RA	Kornblitt A.R., Vibe-Pedersen K., Baralle F.E.;		
RT	"Isolation and Characterization of cDNA clones for human and bovine		
RT	fibronectins.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).		
CC	-!- FUNCTION: Fibronectins bind cell surfaces and various compounds		
CC	including collagen, fibrin, heparin, DNA, and actin. Fibronectins		
CC	are involved in cell adhesion, cell motility, opsonization, wound		
CC	healing, and maintenance of cell shape.		
CC	-!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced		
CC	variants, connected by 2 disulfide bonds near the carboxyl ends;		
CC	to a lesser extend homodimers.		
CC	-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.		
CC	-!- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=1;		
CC	Comment=A number of isoforms are produced. Each of the "extra		
CC	domain" and the connecting strand 3 are present in some forms of		
CC	fibronectin and absent in others;		
CC	Name=1;		
CC	Isold=P07589-1; Sequence=Displayed;		
CC	-!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted		
CC	by hepatocytes. Cellular FN (dimeric or cross-linked multimeric		
CC	forms), made by fibroblasts, epithelial and other cell types, is		
CC	deposited as fibrils in the extracellular matrix.		
CC	-!- PTM: Sulfated (By similarity).		
CC	-!- SIMILARITY: Contains 12 fibronectin type I domains.		
CC	-!- SIMILARITY: Contains 2 fibronectin type II domains.		
CC	-!- SIMILARITY: Contains 15 fibronectin type III domains.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; K00800; AAA30521.2; -		
DR			

DR PIR; A26452; FNBO.
 DR HSP; P08253; 1KS0.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000803; Fibrinctnl.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR003962; FN_III_subd.
 DR Pfam; PF00039; fnl; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 15.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00014; FNTYPEIII.
 DR ProDom; PD000995; FN_Type_II; 2.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00203; FIBRONECTIN_2; 2.
 DR PROSITE; PS00853; FN3; 15.
 KW Acute phase; Alternative splicing; Cell adhesion; Cell shape;
 KW Direct protein sequencing; Glycoprotein; Heparin-binding;
 KW Phosphorylation; Plasma; Pyrrolidone carboxylic acid; Repeat;
 KW Sulfation.
 FT MOD_RES 1 Pyrrolidone carboxylic acid.
 FT DOMAIN 21 241 Fibrin- and heparin-binding 1.
 FT DOMAIN 277 577 Collagen-binding.
 FT DNA_BIND 876 1141
 FT DOMAIN 1236 1509 Cell-attachment.
 FT DOMAIN 1600 1870 Heparin-binding 2.
 FT DOMAIN 1991 2216 Fibrin-binding 2.
 FT DOMAIN 19 59 Fibrinectin type-I 1.
 FT DOMAIN 64 107 Fibrinectin type-I 2.
 FT DOMAIN 108 151 Fibrinectin type-I 3.
 FT DOMAIN 153 197 Fibrinectin type-I 4.
 FT DOMAIN 198 242 Fibrinectin type-I 5.
 FT DOMAIN 275 314 Fibrinectin type-I 6.
 FT DOMAIN 314 373 Fibrinectin type-II 1.
 FT DOMAIN 374 438 Fibrinectin type-II 2.
 FT DOMAIN 437 480 Fibrinectin type-I 7.
 FT DOMAIN 485 527 Fibrinectin type-I 8.
 FT DOMAIN 528 571 Fibrinectin type-I 9.
 FT DOMAIN 576 668 Fibrinectin type-III 1.
 FT DOMAIN 689 778 Fibrinectin type-III 2.
 FT DOMAIN 780 867 Fibrinectin type-III 3.
 FT DOMAIN 877 964 Fibrinectin type-III 4.
 FT DOMAIN 965 1053 Fibrinectin type-III 5.
 FT DOMAIN 1056 1141 Fibrinectin type-III 6.
 FT DOMAIN 1142 1234 Fibrinectin type-III 7.
 FT DOMAIN 1235 1325 Fibrinectin type-III 8.
 FT DOMAIN 1326 1415 Fibrinectin type-III 9.
 FT DOMAIN 1416 1505 Fibrinectin type-III 10.
 FT DOMAIN 1510 1599 Fibrinectin type-III 11 (extra domain).
 FT DOMAIN 1602 1689 Fibrinectin type-III 12.
 FT DOMAIN 1692 1780 Fibrinectin type-III 13.
 FT DOMAIN 1781 1870 Fibrinectin type-III 14.
 FT DOMAIN 1871 1950 Fibrinectin type-III 15.
 FT DOMAIN 1979 2069 Connecting strand 3 (CS-3) (V region).
 FT DOMAIN 2083 2127 Fibrinectin type-III 10.
 FT DOMAIN 2128 2170 Fibrinectin type-I 11.
 FT DOMAIN 2172 2215 Fibrinectin type-I 12.
 FT SITE 1493 1495 Cell attachment site.
 FT DISULFID 21 47
 FT DISULFID 45 56
 FT DISULFID 66 94
 FT DISULFID 92 104
 FT DISULFID 110 138
 FT DISULFID 136 148
 FT DISULFID 155 184
 FT DISULFID 182 194
 FT DISULFID 200 229
 FT DISULFID 227 239
 FT DISULFID 277 304
 FT DISULFID 302 311

FT DISULFID 329 355
 FT DISULFID 343 370
 FT DISULFID 389 415
 FT DISULFID 403 430
 FT DISULFID 439 467
 FT DISULFID 465 477
 FT DISULFID 487 514
 FT DISULFID 512 524
 FT DISULFID 530 558
 FT DISULFID 556 568
 FT DISULFID 2085 2114
 FT DISULFID 2112 2124
 FT DISULFID 2130 2157
 FT DISULFID 2155 2167
 FT DISULFID 2174 2200
 FT DISULFID 2198 2209
 FT DISULFID 2246 2246
 FT DISULFID 2250 2250
 FT MOD_RES 845 845
 FT MOD_RES 850 850
 FT CARBOHYD 399 399
 FT CARBOHYD 497 497
 FT CARBOHYD 511 511
 FT CARBOHYD 846 846
 FT CARBOHYD 976 976
 FT CARBOHYD 1213 1213
 FT CARBOHYD 1387 1387
 FT CARBOHYD 1943 1943
 FT CARBOHYD 1944 1944
 FT MOD_RES 2263 2263
 SQ SEQUENCE 2265 AA; 249557 MW; C2D21D486F498D5C CRC64;
 Query Match 92.7%; Score 255; DB 1; Length 2265;
 Best Local Similarity 89.8%; Pred. No. 1.2e-22;
 Matches 44; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 QAOQWVQSPVAVSQSKPGCYDNGKHVQINQWERTYLGNVLVCTCYG 49
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 1 QAOQVQPSPLTVSQSKPGCYDNGKHVQINQWERTYLGNSALVCTCYG 49
 RESULT 14
 FINE_XENLA STANDARD; PRT; 2481 AA.
 AC Q91740;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Fibronectin precursor.
 GN Name=FN1;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92111942; PubMed=1730390;
 RA Desimone D.W., Norton P.A., Hynes R.O.;
 RT "Identification and characterization of alternatively spliced
 fibronectin mRNAs expressed in early Xenopus embryos.";
 RL Dev. Biol. 149:357-369(1992).
 CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
 including collagen, fibrin, heparin, DNA, and actin. Fibronectins
 are involved in cell adhesion, cell motility, opsonization, wound
 healing, and maintenance of cell shape (By similarity).
 CC -!- SUBUNIT: Dimers or multimers of alternatively spliced variants,
 connected by 2 disulfide bonds near the carboxyl ends (By
 similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=1;
 CC Comment=A number of isoforms are produced. Each of the "extra
 domain" and the connecting strand 3 are present in some forms of

CC fibronectin and absent in others;
 CC Name=1;
 CC IsoId-Q91740-1; Sequence=Displayed;
 CC TISSUE SPECIFICITY: In early Xenopus embryo, cellular forms of
 CC fibronectin predominate which include both extra domains. In
 CC fibronectin of embryonic and adult liver the connecting strand 3
 CC can be either completely excluded or included.
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M77820; AAA49707.1; --
 CC HSP; P08253; IJ7M.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000083; Fibrinctn1.
 CC InterPro; IPR003961; FN III.
 CC InterPro; IPR008957; FN III-like.
 CC InterPro; IPR000562; FN_Type_II.
 CC Pfam; PF00039; fn1; 12.
 CC Pfam; PF00040; fn2; 2.
 CC Pfam; PF00041; fn3; 17.
 CC PRINTS; PD00013; FNTYPEII.
 CC ProDom; PD000395; FN_Type_II; 2.
 CC SMART; SM00058; FN1; 12.
 CC SMART; SM00059; FN2; 2.
 CC SMART; SM00060; FN3; 17.
 CC PROSITE; PS00022; EGF 1; 2.
 CC PROSITE; PS01253; FIBRONECTIN_1; 11.
 CC PROSITE; PS50853; FN3; 17.
 CC Acute phase; Alternative splicing; Cell adhesion; Glycoprotein;
 KW Heparin-binding; Plasma; Repeat; Signal.
 FT SIGNAL 1 31 Potential.
 FT CHAIN 32 2481 Fibronectin.
 FT DOMAIN 55 275 Fibrin- and heparin-binding 1.
 FT DOMAIN 309 609 Fibrin-binding.
 FT DNA_BIND 907 1172 Collagen-binding.
 FT DOMAIN 1358 1631 Cell-attachment.
 FT DOMAIN 1812 2082 Heparin-binding 2.
 FT DOMAIN 2301 2432 Fibrin-binding 2.
 FT DOMAIN 53 93 Fibronectin type-I 1.
 FT DOMAIN 98 141 Fibronectin type-I 2.
 FT DOMAIN 142 185 Fibronectin type-I 3.
 FT DOMAIN 187 231 Fibronectin type-I 4.
 FT DOMAIN 232 276 Fibronectin type-I 5.
 FT DOMAIN 307 346 Fibronectin type-I 6.
 FT DOMAIN 346 405 Fibronectin type-II 1.
 FT DOMAIN 406 470 Fibronectin type-II 2.
 FT DOMAIN 469 512 Fibronectin type-I 7.
 FT DOMAIN 517 559 Fibronectin type-I 8.
 FT DOMAIN 560 603 Fibronectin type-I 9.
 FT DOMAIN 608 700 Fibronectin type-III 1.
 FT DOMAIN 719 809 Fibronectin type-III 2.
 FT DOMAIN 811 898 Fibronectin type-III 3.
 FT DOMAIN 908 995 Fibronectin type-III 4.
 FT DOMAIN 996 1084 Fibronectin type-III 5.
 FT DOMAIN 1087 1172 Fibronectin type-III 6.
 FT DOMAIN 1173 1265 Fibronectin type-III 7.
 FT DOMAIN 1266 1356 Fibronectin type-III 8 (extra domain).
 FT DOMAIN 1357 1447 Fibronectin type-III 9.
 FT DOMAIN 1448 1537 Fibronectin type-III 10.
 FT DOMAIN 1538 1627 Fibronectin type-III 11.
 FT DOMAIN 1632 1721 Fibronectin type-III 12.
 FT DOMAIN 1722 1811 Fibronectin type-III 13 (extra domain).
 FT DOMAIN 1814 1901 Fibronectin type-III 14.
 FT DOMAIN 1904 1992 Fibronectin type-III 15.

FT DOMAIN 1993 2082 Fibronectin type-III 16.
 FT DOMAIN 2083 2205 Connecting strand 3 (CS-3) (V region).
 FT DOMAIN 2192 2284 Fibronectin type-III 17.
 FT DOMAIN 2299 2343 Fibronectin type-I 10.
 FT DOMAIN 2344 2386 Fibronectin type-I 11.
 FT DOMAIN 2388 2431 Fibronectin type-I 12.
 FT SITE 1615 1617 Cell attachment site.
 FT DISULFID 55 81 By similarity.
 FT DISULFID 79 90 By similarity.
 FT DISULFID 100 128 By similarity.
 FT DISULFID 126 138 By similarity.
 FT DISULFID 144 172 By similarity.
 FT DISULFID 170 182 By similarity.
 FT DISULFID 189 218 By similarity.
 FT DISULFID 216 228 By similarity.
 FT DISULFID 234 263 By similarity.
 FT DISULFID 261 273 By similarity.
 FT DISULFID 309 336 By similarity.
 FT DISULFID 334 343 By similarity.
 FT DISULFID 361 387 By similarity.
 FT DISULFID 375 402 By similarity.
 FT DISULFID 421 447 By similarity.
 FT DISULFID 471 499 By similarity.
 FT DISULFID 497 509 By similarity.
 FT DISULFID 519 546 By similarity.
 FT DISULFID 544 556 By similarity.
 FT DISULFID 562 590 By similarity.
 FT DISULFID 588 600 By similarity.
 FT DISULFID 2301 2330 By similarity.
 FT DISULFID 2328 2340 By similarity.
 FT DISULFID 2346 2373 By similarity.
 FT DISULFID 2371 2383 By similarity.
 FT DISULFID 2390 2414 By similarity.
 FT DISULFID 2412 2428 By similarity.
 FT DISULFID 2459 2459 Interchain (with C-2463) (By similarity).
 FT DISULFID 2463 2463 Interchain (with C-2459) (By similarity).
 FT CARBOHYD 431 431 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 529 529 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 543 543 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 877 877 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1244 1244 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1291 1291 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2202 2202 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 2481 AA; 272678 MW; 7847DF4F6CE72C93 CRC64;

Query Match 70.5%; Score 194; DB 1; Length 2481;
 Best Local Similarity 71.4%; Pred. No. 5.5e-15;
 Matches 35; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 QAOQWVQSPVAVSQSPGCGYDNGKHVQINQOQWERTYLGTVLCTCYG 49
 Db 37 QOQVVPQPG--TQDNHKGCGYDNGKYVQINQOQWERTYLGTVLCTCYG 83

RESULT 15
 Q6GQAS PRELIMINARY; PRT; 2481 AA.
 AC Q6GQAS; AC Q6GQAS5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Fibronectin protein.
 GN Name=fibronectin;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCBI_TaxId=8355;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072841; AAH72841.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR00362; FN_Type_II.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 17.
 DR PRINTS; PR00013; FNTYPEII.
 DR ProDom; PD000995; FN_Type_II; 2.
 DR SMART; SM00058; FN1; 12.
 DR SMART; SM00059; FN2; 2.
 DR SMART; SM00060; FN3; 17.
 DR PROSITE; PS00022; EGF 1; UNKNOWN 2.
 DR PROSITE; PS01253; FIBRONECTIN 1; 11.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS50853; FN3; 17.
 DR PROSITE; PS00678; WD REPEATS 1; UNKNOWN 1.
 SQ SEQUENCE 2481 AA; 272708 MW; 4DB47E656DB8B5A3 CRC64;

Query Match 70.5%; Score 194; DB 2; Length 2481;
 Best Local Similarity 71.4%; Pred. No. 5.5e-15;
 Matches 35; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

Qy 1 QAAQMVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGNVLVCTCYG 49
 Db 37 QQQQVVQPQG--TQDNHQKGCYDNGKYYQINQWERTYLGNTLVCTCYG 83

Search completed: September 30, 2005, 13:37:56
 Job time : 36.4502 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 08:57:15 ; Search time 6.30225 Seconds
(without alignments)
580.396 Million cell updates/sec

Title: US-09-492-971b-15_COPY_4_52

Perfect score: 275
Sequence: 1 QAOQWQSPFVAVSQSKPG.....INQWERTYLGWLVCTCYG 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	275	100.0	231	1	US-08-153-799-16
2	275	100.0	224	1	US-08-283-857-1
3	275	100.0	224	5	PCT-US93-09819-1
4	275	100.0	227	6	5455158-1
5	275	100.0	227	6	5455158-1
6	275	100.0	286	2	US-09-016-366A-12
7	275	100.0	246	2	US-08-551-356-2
8	275	100.0	246	5	PCT-US93-12687-2
9	271	98.5	286	4	US-09-961-403-1
10	110	40.0	188	1	US-08-142-449B-14
11	110	40.0	422	2	US-08-836-854-12
12	110	40.0	474	2	US-08-836-854-9
13	80	29.1	42	1	US-08-340-812-7
14	80	29.1	42	1	US-08-459-064B-7
15	80	29.1	42	2	US-08-460-421A-7
16	80	29.1	42	5	PCT-US93-00909-7
17	63	22.9	48	2	US-08-982-597A-20
18	63	22.9	48	3	US-09-136-218-20
19	62	22.5	37	2	US-08-982-597A-17
20	62	22.5	37	3	US-09-136-218-17
21	62	22.5	48	2	US-08-982-597A-22
22	62	22.5	48	3	US-09-136-218-22
23	60	21.8	104	4	US-09-270-767-57611
24	60	21.8	429	4	US-09-270-767-42323
25	60	21.8	519	4	US-09-248-796A-19263
26	58	21.1	448	4	US-09-949-016-9960
27	57.5	20.9	940	2	US-08-938-365-4

Sequence 2, Appli
Sequence 18, Appl
Sequence 18, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 7031, Ap
Sequence 7, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 9, Appli
Sequence 183, App
Sequence 183, App
Sequence 183, App
Sequence 183, App
Sequence 157, App

ALIGNMENTS

RESULT 1
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..2231
; OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match 100.0%; Score 275; DB 1; Length 2231;
Best Local Similarity 100.0%; Pred. No. 4.9e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYG 49
Db 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYG 49

RESULT 2
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,857
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-3528
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US-08-283-857-1

Query Match 100.0%; Score 275; DB 1; Length 2324;
Best Local Similarity 100.0%; Pred. No. 5.1e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYG 49
Db 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYG 49

RESULT 3
PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match 100.0%; Score 275; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 5.1e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYG 49
Db 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYG 49

RESULT 4
5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WEBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
```

```
;SEQ ID NO:1:
5455158-1
    Query Match      100.0%; Score 275; DB 6; Length 2327;
    Best Local Similarity 100.0%; Pred. No. 5.1e-27;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 49
Db 4 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 52

RESULT 5
5455158-1
;PATENT NO. 5455158
;APPLICANT: VOGEL, TIKVA;LEVANON, AVIGDOR;WERBER, MOSHE M.;
;GUY, RACHEL;PANET, AMOS
;TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
;USES AND METHODS OF PRODUCING SAME
;CURRENT APPLICATION DATA:
;NUMBER OF SEQUENCES: 20
;FILING DATE: 04-MAY-1993
;PRIOR APPLICATION NUMBER: US/08/58,241
;APPLICATION NUMBER: 526,397
;FILING DATE: 21-MAY-1990
;APPLICATION NUMBER: 345,952
;FILING DATE: 28-APR-1989
;APPLICATION NUMBER: 291,951
;FILING DATE: 29-DEC-1988
;SEQ ID NO:1:
    LENGTH: 2327
5455158-1

    Query Match      100.0%; Score 275; DB 6; Length 2327;
    Best Local Similarity 100.0%; Pred. No. 5.1e-27;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 49
Db 4 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 52

RESULT 6
US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: NAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997

;SEQ ID NO:1:
5455158-1
    Query Match      100.0%; Score 275; DB 6; Length 2327;
    Best Local Similarity 100.0%; Pred. No. 5.1e-27;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 49
Db 4 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 52

RESULT 7
US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246
; REFERENCE/DOCKET NUMBER: 92-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-356-2

    Query Match      100.0%; Score 275; DB 2; Length 2446;
    Best Local Similarity 100.0%; Pred. No. 5.4e-27;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 49
Db 32 QAOQMVPQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLVCTCYG 80

ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-016-366A-12

    Query Match      100.0%; Score 275; DB 2; Length 2386;
    Best Local Similarity 100.0%; Pred. No. 5.3e-27;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

; GENERAL INFORMATION:


```

RESULT 12
US-08-836-854-9
; Sequence 9, Application US/08836854
; Patent No. 5824547
; GENERAL INFORMATION:
; APPLICANT: HASHINO, Kimikazu
; APPLICANT: MATSUSHITA, Hideyuki
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: METHOD OF PRODUCTION OF TRANSFECTED CELLS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

TELEFAX: 019-333-8943
: INFORMATION FOR SEO ID NO: 7:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-340-812-7

Query Match 29.1%; Score 80; DB 1; Length 42;
Best Local Similarity 42.9%; Pred. No. 0.0017;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 22 YDNGKHVQINQOWERTYLGVLVCTCYG 49
Db 4 YDDGKTYHVGEQWQKEYLGAISSSTSG 31

RESULT 14
US-08-459-064B-7
; Sequence 7, Application US/08459064B
; Patent No. 5747452
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: A METHOD OF MODULATING TUMOR CELL MIGRATION
; TITLE OF INVENTION: USING FIBRONECTIN TYPE III PEPTIDES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,064B
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,462
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; APPLICATION NUMBER: US 08/340,812
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1543
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-459-064B-7

Query Match 29.1%; Score 80; DB 1; Length 42;
Best Local Similarity 42.9%; Pred. No. 0.0017;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 22 YDNGKHVQINQOWERTYLGVLVCTCYG 49
Db 4 YDDGKTYHVGEQWQKEYLGAISSSTSG 31

RESULT 15
US-08-460-421A-7
; Sequence 7, Application US/08460421A
; Patent No. 5837813
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
; MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,421A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,462
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; APPLICATION NUMBER: US 08/340,812
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1542
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-421A-7

Query Match 29.1%; Score 80; DB 2; Length 42;
Best Local Similarity 42.9%; Pred. No. 0.0017;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 22 YDNGKHVQINQOWERTYLGVLVCTCYG 49
Db 4 YDDGKTYHVGEQWQKEYLGAISSSTSG 31

Search completed: September 30, 2005, 08:58:05
Job time : 7.30225 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 08:57:20 ; Search time 37.0257 Seconds
(without alignments)
549.366 Million cell updates/sec

Title: US-09-492-971b-15_COPY_4_52

Perfect score: 275

Sequence: 1 QAQMVPQSPVAUSQKPG.....INQWERTYLGVLVCTCYG 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	275	100.0	259	10 US-09-940-235-4
2	275	100.0	2328	14 US-10-171-311-64
3	275	100.0	2328	15 US-10-236-031B-70
4	275	100.0	2328	15 US-10-374-979-98
5	275	100.0	2328	15 US-10-182-936A-98
6	275	100.0	2328	16 US-10-477-238A-677
7	275	100.0	2328	16 US-10-680-287A-677
8	275	100.0	2328	17 US-10-477-173-677
9	275	100.0	2474	18 US-10-450-763-52638
10	271	98.5	231	9 US-09-925-302-548
11	271	98.5	231	10 US-09-925-302-548

Sequence 4, Appli
Sequence 64, Appl
Sequence 70, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 677, App
Sequence 677, App
Sequence 677, App
Sequence 52638, A
Sequence 548, App
Sequence 548, App

12	271	98.5	463	15	US-10-144-194A-52	Sequence 52, Appl
13	271	98.5	463	16	US-10-491-566-52	Sequence 52, Appl
14	271	98.5	642	16	US-10-741-601-354	Sequence 354, App
15	271	98.5	642	17	US-10-741-600-1066	Sequence 1066, App
16	271	98.5	657	16	US-10-741-601-359	Sequence 359, App
17	271	98.5	657	17	US-10-741-600-1072	Sequence 1072, App
18	271	98.5	984	16	US-10-741-601-356	Sequence 356, App
19	271	98.5	984	17	US-10-741-600-1069	Sequence 1069, App
20	271	98.5	1173	18	US-10-450-763-52634	Sequence 52634, A
21	271	98.5	2220	15	US-10-236-392-4	Sequence 4, Appli
22	271	98.5	2296	16	US-10-741-601-363	Sequence 363, App
23	271	98.5	2296	17	US-10-741-600-1075	Sequence 1075, App
24	271	98.5	2355	15	US-10-144-194A-104	Sequence 104, App
25	271	98.5	2355	15	US-10-360-101-235	Sequence 235, App
26	271	98.5	2355	15	US-10-447-161-3	Sequence 3, Appli
27	271	98.5	2355	16	US-10-734-564-94	Sequence 94, Appl
28	271	98.5	2355	16	US-10-741-601-357	Sequence 357, App
29	271	98.5	2355	16	US-10-741-601-366	Sequence 366, App
30	271	98.5	2355	16	US-10-491-566-104	Sequence 104, App
31	271	98.5	2355	17	US-10-741-600-1067	Sequence 1067, App
32	271	98.5	2355	17	US-10-741-600-1078	Sequence 1078, App
33	271	98.5	2355	17	US-10-852-335A-147	Sequence 147, App
34	271	98.5	2355	18	US-10-287-436A-436	Sequence 436, App
35	271	98.5	2355	18	US-10-287-436A-1137	Sequence 1137, App
36	271	98.5	2355	20	US-11-040-130-28	Sequence 28, Appl
37	271	98.5	2386	10	US-09-961-403-1	Sequence 1, Appli
38	271	98.5	2386	16	US-10-741-601-360	Sequence 360, App
39	271	98.5	2386	16	US-10-788-792-206	Sequence 206, App
40	271	98.5	2386	16	US-10-618-281-32	Sequence 32, Appl
41	271	98.5	2386	17	US-10-741-600-1071	Sequence 1071, App
42	271	98.5	2386	17	US-10-868-577A-59	Sequence 59, Appl
43	271	98.5	2386	18	US-10-485-758-4	Sequence 4, Appli
44	271	98.5	2386	18	US-10-485-758-9	Sequence 9, Appli
45	256	93.1	2320	15	US-10-236-392-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-940-235-4
; Sequence 4, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT FILING DATE: 2002-04-09
; PRIOR FILING DATE: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 259
; ORGANISM: Homo sapiens
US-09-940-235-4

Query Match 100.0%; Score 275; DB 10; Length 259;
Best Local Similarity 100.0%; Pred. No. 5.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
/ ORGANISM: Homo sapiens
US-10-182-936A-98

Query Match      100.0%; Score 275; DB 15; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.3e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 49
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 53

RESULT 6
US-10-477-238A-677
; Sequence 677, Application US/10477238A
; Publication No. US20040221326A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-238A-677

Query Match      100.0%; Score 275; DB 16; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.3e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 49
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 53

RESULT 7
US-10-680-287A-677
; Sequence 677, Application US/10680287A
; Publication No. US2004024069A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
```

```
/ NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-680-287A-677

Query Match      100.0%; Score 275; DB 16; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.3e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 49
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 53

RESULT 8
US-10-477-173-677
; Sequence 677, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-173-677

Query Match      100.0%; Score 275; DB 17; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.3e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 49
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYINQOQWERTYLGNVLVCTCYG 53

RESULT 9
US-10-450-763-52638
; Sequence 52638, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
```

; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 52638
; LENGTH: 2474
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (504)..(541)
; OTHER INFORMATION: Type II fibronectin collagen-binding domain proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00023, p-value=8.920e-3
; OTHER INFORMATION: raw score of 24.31
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (701)..(2267)
; OTHER INFORMATION: Fibronectin type III domain identified by Pfam, accession
; OTHER INFORMATION: name fn3, E-value=3.1e-275, Pfam score of 927.8
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2474)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
; US-10-450-763-52638

Query Match 100.0%; Score 275; DB 18; Length 2474;
Best Local Similarity 100.0%; Pred. No. 6.8e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 49
Db 123 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 171

RESULT 10
US-09-925-302-548
; Sequence 548, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-548

Query Match 98.5%; Score 271; DB 9; Length 231;
Best Local Similarity 98.0%; Pred. No. 1.6e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 49
Db 68 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 116

RESULT 11
US-09-925-302-548
; Sequence 548, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-548

Query Match 98.5%; Score 271; DB 15; Length 463;
Best Local Similarity 98.0%; Pred. No. 3.4e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 49
Db 68 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 116

RESULT 12
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-194A-52

Query Match 98.5%; Score 271; DB 15; Length 463;
Best Local Similarity 98.0%; Pred. No. 3.4e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 49
Db 68 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 116

Db 61 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 109

RESULT 11
US-09-925-302-548
; Sequence 548, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-548

Query Match 98.5%; Score 271; DB 10; Length 231;
Best Local Similarity 98.0%; Pred. No. 1.6e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 49
Db 61 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 109

RESULT 12
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-194A-52

Query Match 98.5%; Score 271; DB 15; Length 463;
Best Local Similarity 98.0%; Pred. No. 3.4e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 49
Db 68 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOQWERTYLGNAVLCVCTCYG 116

```
RESULT 13
US-10-491-566-52
; Sequence 52, Application US/10491566
; Publication No. US20040249144A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/491,566
; CURRENT FILING DATE: 2004-04-05
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-491-566-52

Query Match          98.5%; Score 271; DB 16; Length 463;
Best Local Similarity 98.0%; Pred. No. 3.4e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOQWERTYLGNNLVCTCYG 49
Db 68 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOQWERTYLGNNLVCTCYG 116

RESULT 14
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match          98.5%; Score 271; DB 16; Length 642;
Best Local Similarity 98.0%; Pred. No. 4.9e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOQWERTYLGNNLVCTCYG 49
Db 32 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOQWERTYLGNNLVCTCYG 80

RESULT 15
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match          98.5%; Score 271; DB 17; Length 642;
Best Local Similarity 98.0%; Pred. No. 4.9e-26;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOQWERTYLGNNLVCTCYG 49
Db 32 QAOQMVQPSPVAVSQSKPGCYDNGKHQYQINQOQWERTYLGNNLVCTCYG 80

Search completed: September 30, 2005, 09:21:21
Job time : 37.0257 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: September 30, 2005, 13:30:52 ; Search time 189.55 Seconds
(without alignments)
707.807 Million cell updates/sec

Title: US-09-492-971b-15_COPY_4_265

Perfect score: 1499

Sequence: 1 QAAQWVQSPVAVSQSKPG.....TSVQTSSGSPFTDRAAV 262

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1495	99.7	749	Q72391	homo sapien
2	1495	99.7	1103	Q6MZF4	homo sapien
3	1495	99.7	2296	Q6NOA6	homo sapien
4	1495	99.7	2386	1 FINC HUMAN	homo sapien
5	1495	99.7	2477	Q6MZU5	homo sapien
6	1494	99.7	2357	Q6BDT4	homo sapien
7	1491	99.5	2267	Q6BDP9	homo sapien
8	1491	99.5	2444	Q6NO25	homo sapien
9	1474	98.3	2240	Q6BDP8	homo sapien
10	1447	96.5	2265	1 FINC BOVIN	homo sapien
11	1419.5	94.7	2477	1 FINC MOUSE	mus musculus
12	1412.5	94.2	2477	1 FINC RAT	rattus norvegicus
13	1382.5	92.2	236	Q8C6J7	mus musculus
14	1296	86.5	2491	1 FINC XENLA	xenopus laevis
15	1296	86.5	2491	Q6GQA5	xenopus laevis
16	1103.5	73.6	922	Q93405	brachydanio
17	1103.5	73.6	2478	Q93406	brachydanio
18	1042.5	69.5	2408	Q6JAN2	brachydanio
19	507	33.8	215	Q6DD34	xenopus laevis
20	362.5	24.2	2193	Q6MZW7	homo sapien
21	298	19.9	412	Q71U44	homo sapien
22	296.5	19.8	216	Q99KD0	mus musculus
23	296.5	19.8	810	Q8R3F3	mus musculus
24	294.5	19.6	211	Q7L553	homo sapien
25	294.5	19.6	268	Q6PJES	homo sapien
26	294.5	19.6	1255	Q6MZS0	homo sapien
27	284.5	19.6	2146	Q68CX6	homo sapien
28	287.5	19.2	141	Q90XQ2	ambystoma m
29	287.5	19.2	1034	Q6N084	homo sapien
30	285.5	19.0	351	Q02816	oryctolagus
31	278.5	18.6	358	Q7T1S2	brachydanio

RESULT 1

```
Q72391 ID Q72391 PRELIMINARY; PRT; 749 AA.
AC Q72391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
DE Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBT_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -
DR HSP; Q96KPF; 1FBR.
DR GO; GO:0005576; C:extracellular; IRA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibronctnl.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;
```

Query Match 99.7%; Score 1495; DB 2; Length 749;

Best Local Similarity 99.6%; Pred. No. 3.1e-121;

Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 QAAQWVQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 60
DB 124 QAAQWVQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 183
QY 61 EAETCTFDKYTGNTYRVGDTYRPRKDSMTWDCCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 184 EAETCTFDKYTGNTYRVGDTYRPRKDSMTWDCCTCIGAGRGRICTIANRCHGGQSYKIG 243
QY 121 DTWRPHETGGYMLCVCVLGNGKGWETCKPIAEKCPDHAAGTSYVVGETWEKPYQGMVMV 180
DB 244 DTWRPHETGGYMLCVCVLGNGKGWETCKPIAEKCPDHAAGTSYVVGETWEKPYQGMVMV 303
QY 181 DCTCLGEGSGRICTSRNRCNDODTFTSYRIGDTWSKDNRNGLLCICTGNRGRGWKCE 240
```

ALIGNMENTS

Db 304 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNRGWEKCE 363
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 364 RHTSVQTTSSGSGPFTDVRAAV 385

RESULT 2
Q6MZFA PRELIMINARY; PRT; 1103 AA.
AC Q6MZFA; 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Hypothetical protein DKFZp666K139 (Hypothetical protein DKFZp666F?19)
DE (Fragment).
GN Name=DKFZp666K139; Synonyms=DKFZp666F219;

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;

RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansorge W., Krieger S., Regiert T., Rittmuller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Oeanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX649182; CAE46200.1; -.
DR EMBL; BX640802; CAE45885.1; -.
DR GO; GO:0005576; C:extracellular; IEA.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR000083; Fibnctnl.

DR InterPro; IPR003961; FN_III.

DR InterPro; IPR008957; FN_III-like.

DR InterPro; IPR000562; FN_Type_II.

DR Pfam; PF00039; fnl; 9.

DR Pfam; PF00040; fn2; 2.

DR Pfam; PF00041; fn3; 4.

DR PRINTS; PR00012; FNTYPEI.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 2.

DR SMART; SM00058; FN1; 9.

DR SMART; SM00059; FN2; 2.

DR SMART; SM00060; FN3; 4.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01253; FIBRONECTIN_1; 9.

DR PROSITE; PS00023; FIBRONECTIN_2; 2.

DR PROSITE; PS00853; FN3; 4.

DR Hypothetical protein.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;

Query Match 99.7%; Score 1495; DB 2; Length 1103;

Best Local Similarity 99.6%; Pred. No. 4.6e-121;

Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAQWVQPSQSPVAVSQSKPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCEKSP 60
Db 121 QAQWVQPSQSPVAVSQSKPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCEKSP 180

QY 61 EAEETCFDKYTGNTYRVGDYTERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 181 EAEETCFDKYTGNTYRVGDYTERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 240

QY 121 DWRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
Db 241 DWRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 300

QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNRGWEKCE 240
Db 301 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNRGWEKCE 360

QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 361 RHTSVQTTSSGSGPFTDVRAAV 382

RESULT 3

Q6NOA6

ID Q6NOA6 PRELIMINARY; PRT; 2296 AA.

AC Q6NOA6;

DT 05-JUL-2004 (TRENBLrel. 27, Created)

DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)

DE Hypothetical protein DKFZp686M04163.

GN Name=DKFZp686M04163;

OS Homo sapiens (Human);

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=Human uterus endothel primary cell culture;

RG The German Human cDNA Consortium;

RA Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Oeanger A.,

RA Fobo G., Han M., Wiemann S.;

RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX640608; CAE45714.1; -.
DR GO; GO:0005576; C:extracellular; IEA.

DR InterPro; IPR002086; Aldehyd dehydrog.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR000083; Fibnctnl.

DR InterPro; IPR003961; FN_III.

DR InterPro; IPR008957; FN_III-like.

DR InterPro; IPR000562; FN_Type_II.

DR Pfam; PF00039; fnl; 12.

DR Pfam; PF00040; fn2; 2.

DR Pfam; PF00041; fn3; 15.

DR PRINTS; PR00012; FNTYPEI.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 2.

DR SMART; SM00058; FN1; 12.

DR SMART; SM00059; FN2; 2.

DR SMART; SM00060; FN3; 15.

DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.

DR PROSITE; PS01253; FIBRONECTIN_1; 12.

DR PROSITE; PS00023; FIBRONECTIN_2; 2.

DR PROSITE; PS00853; FN3; 15.

DR Hypothetical protein.

KW Hypothetical protein.

SQ SEQUENCE 2296 AA; 252761 MW; 9AB2D723CC0CED70 CRC64;

Query Match 99.7%; Score 1495; DB 2; Length 2296;

Best Local Similarity 99.6%; Pred. No. 9.5e-121;

Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAQWVQPSQSPVAVSQSKPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCEKSP 60
Db 32 QAQWVQPSQSPVAVSQSKPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCEKSP 91

QY 61 EAEETCFDKYTGNTYRVGDYTERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 92 EAEETCFDKYTGNTYRVGDYTERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151

QY 121 DWRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
Db 152 DWRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 211

QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNRGWEKCE 240
Db 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNRGWEKCE 271

QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 4
 F1NC HUMAN
 ID F1NC HUMAN STANDARD; PRT; 2386 AA.
 AC P02751; O95609; Q14312; Q14325; Q14326; Q86T27; Q8IV18;
 AC Q96KP7; Q96KP8; Q96KP9; Q9H1B8; Q9HAP3; Q9UMK2;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Fibronectin precursor (FN) (Cold-insoluble globulin) (CIG).
 GN Name=FN1; Synonyms=FN;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21600194; PubMed=11737888; DOI=10.1186/bcr325;
 RA Schor S.L., Schor A.M.;
 RA "Phenotypic and genetic alterations in mammary stroma: implications for tumour progression.";
 RT Breast Cancer Res. 3:373-379(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 3; 7 AND 10).
 RC TISSUE=Cervix;
 RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
 RA Newes H.-W., Weil B., Amid C., Osanger A., Fobo G., Han M.,
 RA Wiemann S.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-38 FROM N.A.
 RX MEDLINE=87030890; PubMed=3770189; DOI=10.1016/0014-5793(86)80029-1;
 RA Gutman A., Yamada K.M., Kornblitt A.R.;
 RT "Human fibronectin is synthesized as a pre-propolypeptide.";
 RL FEBS Lett. 207:145-148(1986).
 RN [4]
 RP SEQUENCE OF 1-49 FROM N.A.
 RX MEDLINE=87175578; PubMed=3031656;
 RA Dean D.C., Bowls C.L., Bourgeois S.;
 RT "Cloning and analysis of the promoter region of the human fibronectin gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:1876-1880(1987).
 RN [5]
 RP SEQUENCE OF 28-2386 FROM N.A. (ISOFORM 3).
 RX MEDLINE=85284965; PubMed=2992939;
 RA Kornblitt A.R., Umezawa K., Vibe-Pedersen K., Baralle F.E.;
 RT "Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene.";
 RL EMBO J. 4:1755-1759(1985).
 RN [6]
 RP SEQUENCE OF 103-481 AND 2228-2386 FROM N.A. (ISOFORMS 1; 3; 8 AND 9).
 RC TISSUE=Peripheral blood T-cell, and Umbilical vein endothelial cells;
 RA Godfrey H.P., Ebrahim A.A.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 973-2386 FROM N.A. (ISOFORM 3).
 RX MEDLINE=84272258; PubMed=6462919;
 RA Kornblitt A.R., Vibe-Pedersen K., Baralle F.E.;
 RT "Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats.";
 RL Nucleic Acids Res. 12:5853-5868(1984).
 RN [8]
 RP SEQUENCE OF 1232-1782 FROM N.A. (ISOFORM 7).
 RX MEDLINE=88233940; PubMed=3375063;
 RA Paoletta G., Henschliffe C., Sebastio G., Baralle F.E.;
 RT "Sequence analysis and in vivo expression show that alternative splicing of ED-B and ED-A regions of the human fibronectin gene are independent events.";
 RL Nucleic Acids Res. 16:3545-3557(1988).
 RN [9]
 RP SEQUENCE OF 1257-1365 FROM N.A. (ISOFORM 11).
 RX MEDLINE=88041070; PubMed=3478690;

RA Gutman A., Kornblitt A.R.;
 RT "Identification of a third region of cell-specific alternative splicing in human fibronectin mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7179-7182(1987).
 RN [10]
 RP SEQUENCE OF 1441-1548.
 RX MEDLINE=82265604; PubMed=7050098;
 RA Pierschbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;
 RT "The cell attachment domain of fibronectin. Determination of the primary structure.";
 RL J. Biol. Chem. 257:9593-9597(1982).
 RN [11]
 RP SEQUENCE OF 1448-1540 FROM N.A.
 RX MEDLINE=83290929; PubMed=6688418;
 RA Oldberg A., Linney E., Ruoslahti E.;
 RT "Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell attachment domain in human fibronectin.";
 RL J. Biol. Chem. 258:10193-10196(1983).
 RN [12]
 RP SEQUENCE OF 1448-1540 FROM N.A.
 RX MEDLINE=86111901; PubMed=3003095;
 RA Oldberg A., Ruoslahti E.;
 RT "Evolution of the fibronectin gene. Exon structure of cell attachment domain.";
 RL J. Biol. Chem. 261:2113-2116(1986).
 RN [13]
 RP SEQUENCE OF 1594-2386 FROM N.A. (ISOFORM 1).
 RX MEDLINE=85280409; PubMed=2992573;
 RA Bernard M.P., Kolbe M., Weil D., Chu M.-L.;
 RT "Human cellular fibronectin: comparison of the carboxyl-terminal portion with rat identifies primary structural domains separated by hypervariable regions.";
 RL Biochemistry 24:2698-2704(1985).
 RN [14]
 RP SEQUENCE OF 1712-1739 FROM N.A.
 RX MEDLINE=87026578; PubMed=3021206;
 RA Sekiguchi K., Klos A.M., Kurachi K., Yoshitake S., Hakomori S.;
 RT "Human liver fibronectin complementary DNAs: identification of two different messenger RNAs possibly encoding the alpha and beta subunits of plasma fibronectin.";
 RL Biochemistry 25:4936-4941(1986).
 RN [15]
 RP SEQUENCE OF 1788-2386 FROM N.A. (ISOFORMS 4; 5 AND 6).
 RC TISSUE=Cartilage;
 RX MEDLINE=22126816; PubMed=12127832; DOI=10.1053/joca.2002.0792;
 RA Parker A.E., Boutell J., Carr A., Maciewicz R.A.;
 RT "Novel cartilage-specific splice variants of fibronectin.";
 RL Osteoarthritis Cartilage 10:528-534(2002).
 RN [16]
 RP SEQUENCE OF 32-290.
 RX MEDLINE=84032463; PubMed=6630202;
 RA Garcia-Pardo A., Pearlstein E., Frangione B.;
 RT "Primary structure of human plasma fibronectin. The 29,000-dalton NH2-terminal domain.";
 RL J. Biol. Chem. 258:12670-12674(1983).
 RN [17]
 RP SEQUENCE OF 309-608, AND COLLAGEN-BINDING.
 RX MEDLINE=87080265; PubMed=3024962;
 RA Owens R.J., Baralle F.E.;
 RT "Mapping the collagen-binding site of human fibronectin by expression in Escherichia coli.";
 RL EMBO J. 5:2825-2830(1986).
 RN [18]
 RP SULFATION.
 RX MEDLINE=86042625; PubMed=2414772;
 RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
 RT "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
 RN [19]
 RP O-GLYCOSYLATION OF THR-2064.
 RX MEDLINE=91190085; PubMed=2012601;
 RA Tresselt T., McCarthy J.B., Calaycay J., Lee T.D., Legesse K.,

RA Shively J.E., Pande H.;
 RT "Human plasma fibronectin. Demonstration of structural differences
 between the A- and B-chains in the III CS region.";
 RL Biochem. J. 274:731-738(1991).
 [20]
 RP FELM1-BINDING SITE.
 RX MEDLINE=93015879; PubMed=1400330;
 RA Balbona K., Tran H., Godyna S., Ingham K.C., Strickland D.K.,
 RA Argaves W.S.;
 RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding
 region of fibronectin.";
 RL J. Biol. Chem. 267:20120-20125(1992).
 [21]
 RN CHARACTERIZATION OF FIBRIN-BINDING SITE 1.
 RX MEDLINE=95081153; PubMed=7989369;
 RA Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;
 RT "Further characterization of the NH2-terminal fibrin-binding site on
 fibronectin.";
 RL J. Biol. Chem. 269:31938-31945(1994).
 [22]
 RN INTERACTION WITH LGALS3BP.
 RX PubMed=9501082; DOI=10.1093/emboj/17.6.1606;
 RA Sasaki T., Brakebusch C., Engel J., Timpl R.;
 RT "Mac-2 binding protein is a cell-adhesive protein of the extracellular
 matrix which self-assembles into ring-like structures and binds beta1
 integrins, collagens and fibronectin.";
 RL EMBO J. 17:1606-1613(1998).
 [23]
 RN STRUCTURE BY NMR OF 1447-1540.
 RX MEDLINE=92162710; PubMed=1311120;
 RA Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,
 RA Campbell I.D.;
 RT "1H NMR assignment and secondary structure of the cell adhesion type
 III module of fibronectin.";
 RL Biochemistry 31:2068-2073(1992).
 [24]
 RN STRUCTURE BY NMR OF 1447-1540.
 RX MEDLINE=93046665; PubMed=1423622; DOI=10.1016/0092-8674(92)90600-H;
 RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;
 RT "The three-dimensional structure of the tenth type III module of
 fibronectin: an insight into RGD-mediated interactions.";
 RL Cell 71:671-678(1992).
 [25]
 RN STRUCTURE BY NMR OF 182-275.
 RX MEDLINE=94141923; PubMed=8308892;
 RA Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,
 RA Campbell I.D.;
 RT "Solution structure of a pair of fibronectin type 1 modules with
 fibrin binding activity.";
 RL J. Mol. Biol. 235:1302-1311(1994).
 [26]
 RN STRUCTURE BY NMR OF 32-92.
 RX MEDLINE=96069779; PubMed=7583666;
 RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;
 RT "High-resolution structural studies of the factor XIIIa crosslinking
 site and the first type 1 module of fibronectin.";
 RL Nat. Struct. Biol. 2:946-950(1995).
 [27]

Query Match 99.7%; Score 1495; DB 1; Length 2386;
 Best Local Similarity 99.6%; Pred. No. 9.9e-121;
 Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAAQMVQPSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 60
 DB 32 QAAQMVQPSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 91
 QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTIANRCHGGQSYKIG 120
 DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTIANRCHGGQSYKIG 151
 QY 121 DTWRRPHETGGYMLCVCGLGKNGKGTCKPIAEKCFDHAAGTSYVVGTEWKEPYQGMMV 180

DB 152 DTWRRPHETGGYMLCVCGLGKNGKGTCKPIAEKCFDHAAGTSYVVGTEWKEPYQGMMV 211
 QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNGRGWKCE 240
 DB 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNGRGWKCE 271
 QY 241 RHTSVQTTSSGSGPFTDVRRAV 262
 DB 272 RHTSVQTTSSGSGPFTDVRRAV 293
 RESULT 5
 Q6MZU5
 ID Q6MZU5 PRELIMINARY; PRT; 2477 AA.
 AC Q6MZU5;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Hypothetical protein DKFP68601166.
 GN Name=DKFP68601166;
 OS Homo sapiens (Human),
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Human endometrium carcinoma cell line;
 RG The German Human CDNA Consortium;
 RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
 RA Fobo G., Han M., Wiemann S.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX640875; CAE45932.1;
 DR GO; GO:0005576; C:extracellular; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR000562; FN_Type_II.
 DR Pfam; PF00039; fnl; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 17.
 DR PRINTS; PR00013; FNTYPEII.
 DR ProDom; PD000995; FN_Type_II; 2.
 DR SMART; SM00058; FN1; 12.
 DR SMART; SM00059; FN2; 2.
 DR SMART; SM00060; FN3; 17.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.
 DR PROSITE; PS00853; FN3; 17.
 KW Hypothetical protein.
 SQ SEQUENCE 2477 AA; 272335 MW; D358D85C6B18207C CRC64;

Query Match 99.7%; Score 1495; DB 2; Length 2477;
 Best Local Similarity 99.6%; Pred. No. 1e-120;
 Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAAQMVQPSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 60
 DB 32 QAAQMVQPSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 91
 QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTIANRCHGGQSYKIG 120
 DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTIANRCHGGQSYKIG 151
 QY 121 DTWRRPHETGGYMLCVCGLGKNGKGTCKPIAEKCFDHAAGTSYVVGTEWKEPYQGMMV 180
 DB 152 DTWRRPHETGGYMLCVCGLGKNGKGTCKPIAEKCFDHAAGTSYVVGTEWKEPYQGMMV 211
 QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKKNRGNLLQCICTGNGRGWKCE 240

```

Db      212 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTWSKKNRGNLLQICITGNRGEWKCE 271
QY      241 RHTSVQTTSSSGSPFTDVRAAV 262
Db      272 RHTSVQTTSSSGSPFTDVRAAV 293

RESULT 6
Q68DT4
ID      Q68DT4      PRELIMINARY;      PRT;      2357 AA.
AC      Q68DT4;
DT      25-OCT-2004 (TReMBLrel. 28, Created)
DT      25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE      Hypothetical protein DKFZp686F10164.
GN      Name=DKFZp686F10164;
OS      Homo sapiens (Human);
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RG      TISSUE=Uterus endothel;
RA      Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA      Fobo G., Han M., Wiemann S.;
RL      Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; CR749281; CAH18136.1; -.
DR      InterPro; IPR002086; Aldehyd dehydrog.
DR      InterPro; IPR006209; EGF like.
DR      InterPro; IPR000083; Fibrnctnl.
DR      InterPro; IPR003962; FNIII subd.
DR      InterPro; IPR003961; FNIII.
DR      InterPro; IPR008957; FNIII-like.
DR      InterPro; IPR000562; FN_Type_II.
DR      Pfam; PF00039; fn1; 12.
DR      Pfam; PF00040; fn2; 2.
DR      Pfam; PF00041; fn3; 15.
DR      PRINTS; PR00012; FNTYPEI.
DR      PRINTS; PR00013; FNTYPEII.
DR      PRINTS; PR00014; FNTYPEIII.
DR      ProDom; PD000995; FN_Type_II; 2.
DR      SMART; SM00058; FN1; 12.
DR      SMART; SM00059; FN2; 2.
DR      SMART; SM00060; FN3; 15.
DR      PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR      PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR      PROSITE; PS01253; FIBRONECTIN_1; 12.
DR      PROSITE; PS00023; FIBRONECTIN_2; 2.
KW      Hypothetical protein.
SQ      SEQUENCE 2357 AA; 259090 MW; BEAE3990E27E532A CRC64;

Query Match      99.7%; Score 1494; DB 2; Length 2357;
Best Local Similarity 99.2%; Pred. No. 1.2e-120;
Matches 260; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1      QAAQVQPSQSPVAVSQSGPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 60
Db      124      QAAQVQPSQSPVAVSQSGPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 183
QY      61      EAEETCFDKYTGNTYRVGDTYRPKDSMIWDCCTCIGAGGRISCTIANRCHGGQSYKIG 120
Db      184      EAEETCFDKYTGNTYRVGDTYRPKDSMIWDCCTCIGAGGRISCTIANRCHGGQSYKIG 243
QY      121      DTWRPHETGGYMLCEVCLGNKGKGTCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
Db      244      DTWRPHETGGYMLCEVCLGNKGKGTCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 303
QY      181      DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTWSKKNRGNLLQICITGNRGEWKCE 240
Db      304      DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTWSKKNRGNLLQICITGNRGEWKCE 363
QY      241      RHTSVQTTSSSGSPFTDVRAAV 262

```

```

Db      364 RHTSVQTTSSSGSPFTDVRAAV 385

RESULT 7
Q68DP9
ID      Q68DP9      PRELIMINARY;      PRT;      2267 AA.
AC      Q68DP9;
DT      25-OCT-2004 (TReMBLrel. 28, Created)
DT      25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE      Hypothetical protein DKFZp686K08164.
GN      Name=DKFZp686K08164;
OS      Homo sapiens (Human);
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RG      TISSUE=Uterus endothel;
RA      Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA      Fobo G., Han M., Wiemann S.;
RL      Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; CR749316; CAH18171.1; -.
DR      InterPro; IPR002086; Aldehyd dehydrog.
DR      InterPro; IPR006209; EGF like.
DR      InterPro; IPR000083; Fibrnctnl.
DR      InterPro; IPR003962; FNIII subd.
DR      InterPro; IPR003961; FNIII.
DR      InterPro; IPR008957; FNIII-like.
DR      InterPro; IPR000562; FN_Type_II.
DR      Pfam; PF00039; fn1; 12.
DR      Pfam; PF00040; fn2; 2.
DR      Pfam; PF00041; fn3; 16.
DR      PRINTS; PR00012; FNTYPEI.
DR      PRINTS; PR00013; FNTYPEII.
DR      PRINTS; PR00014; FNTYPEIII.
DR      ProDom; PD000995; FN_Type_II; 2.
DR      SMART; SM00058; FN1; 12.
DR      SMART; SM00059; FN2; 2.
DR      SMART; SM00060; FN3; 16.
DR      PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR      PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR      PROSITE; PS01253; FIBRONECTIN_1; 12.
DR      PROSITE; PS00023; FIBRONECTIN_2; 2.
KW      Hypothetical protein.
SQ      SEQUENCE 2267 AA; 249358 MW; C4D124A038C323DF CRC64;

Query Match      99.5%; Score 1491; DB 2; Length 2267;
Best Local Similarity 99.2%; Pred. No. 2.1e-120;
Matches 260; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1      QAAQVQPSQSPVAVSQSGPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 60
Db      32      QAAQVQPSQSPVAVSQSGPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 91
QY      61      EAEETCFDKYTGNTYRVGDTYRPKDSMIWDCCTCIGAGGRISCTIANRCHGGQSYKIG 120
Db      92      EAEETCFDKYTGNTYRVGDTYRPKDSMIWDCCTCIGAGGRISCTIANRCHGGQSYKIG 151
QY      121      DTWRPHETGGYMLCEVCLGNKGKGTCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
Db      152      DTWRPHETGGYMLCEVCLGNKGKGTCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 211
QY      181      DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTWSKKNRGNLLQICITGNRGEWKCE 240
Db      212      DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTWSKKNRGNLLQICITGNRGEWKCE 271
QY      241      RHTSVQTTSSSGSPFTDVRAAV 262
Db      272      RHTSVQTTSSSGSPFTDVRAAV 293

```

```
RESULT 8
Q6N025
ID Q6N025 PRELIMINARY; PRT; 2444 AA.
AC Q6N025;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M2451 (Fragment).
GN Name=DKFZp686M2451;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RG The German Human CDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640733; CAE45847.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 16.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 16.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 16.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2444 AA; 268676 MW; 71C5E8C56A84C7BC CRC64;

Query Match 99.5%; Score 1491; DB 2; Length 2444;
Best Local Similarity 99.2%; Pred. No. 2.3e-120;
Matches 260; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QAOQVQPSFVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGNCEKSP 60
DB 121 QAOQVQPSFVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGNCEKSP 180
QY 61 EAEETCFDKYGTNTYRVGDTYERPKDSMIWDCITCIGAGRGISCTIANRCHGGQSYKIG 120
DB 181 EAEETCFDKYGTNTYRVGDTYERPKDSMIWDCITCIGAGRGISCTIANRCHGGQSYKIG 240
QY 121 DWRPRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGWMMV 180
DB 241 DWRPRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGWMMV 300
QY 181 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDTSWKKDNKGNLLQCTCTGNRGWEKCE 240
DB 301 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDTSWKKDNKGNLLQCTCTGNRGWEKCE 360
QY 241 RHTSVQTTSSSGSGPFTDRAAV 262
DB 361 RHTSVQTTSSSGSGPFTDRAAV 382

RESULT 9
Q68DP8
ID Q68DP8 PRELIMINARY; PRT; 2240 AA.
AC Q68DP8;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE Hypothetical protein DKFZp686H0342.
GN Name=DKFZp686H0342;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RG The German CDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF 1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 2240 AA; 246667 MW; 8FCDAP406F330621 CRC64;

Query Match 98.3%; Score 1474; DB 2; Length 2240;
Best Local Similarity 98.9%; Pred. No. 6.2e-119;
Matches 259; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAOQVQPSFVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGNCEKSP 60
DB 32 QAOQVQPSFVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGNCEKSP 91
QY 61 EAEETCFDKYGTNTYRVGDTYERPKDSMIWDCITCIGAGRGISCTIANRCHGGQSYKIG 120
DB 92 EAEETCFDKYGTNTYRVGDTYERPKDSMIWDCITCIGAGRGISCTIANRCHGGQSYKIG 151
QY 121 DWRPRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGWMMV 180
DB 152 DWRPRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGWMMV 211
QY 181 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDTSWKKDNKGNLLQCTCTGNRGWEKCE 240
DB 212 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDTSWKKDNKGNLLQCTCTGNRGWEKCE 271
QY 241 RHTSVQTTSSSGSGPFTDRAAV 262
DB 272 RHTSVQTTSSSGSGPFTDRAAV 293

RESULT 10
FINC_BOVIN
ID FINC_BOVIN STANDARD; PRT; 2265 AA.
AC P07589;
DT 01-APR-1988 (Rel. 07, Created)
```

DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Fibronection (FN).
 GN Name=FN1;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_taxid=9913;
 RN [1]
 RN SEQUENCE.
 RX MEDLINE=87054047; PubMed=3780752;
 RA Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
 RT "Complete primary structure of bovine plasma fibronectin.";
 RL Eur. J. Biochem. 161:441-453 (1986).
 RN [2]
 RN PARTIAL SEQUENCE.
 RX MEDLINE=83117805; PubMed=6218503;
 RA Petersen T.E., Thorgersen H.C., Skorstengaard K., Vibe-Pedersen K.,
 RA Sahl P., Sottrup-Jensen L., Magnusson S.;
 RT "Partial primary structure of bovine plasma fibronectin: three types
 of internal homology";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141 (1983).
 RN [3]
 RN SEQUENCE OF 2170-2265 FROM N.A.
 RX MEDLINE=83221567; PubMed=6304699;
 RA Kornblith A.R., Vibe-Pedersen K., Baralle F.E.;
 RT "Isolation and characterization of cDNA clones for human and bovine
 fibronectins";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222 (1983).
 CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
 including collagen, fibrin, heparin, DNA, and actin. Fibronectins
 are involved in cell adhesion, cell motility, opsonization, wound
 healing, and maintenance of cell shape.
 CC -!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
 variants, connected by 2 disulfide bonds near the carboxyl ends;
 to a lesser extend homodimers.
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=1;
 CC Comment=A number of isoforms are produced. Each of the "extra
 domain" and the connecting strand 3 are present in some forms of
 fibronectin and absent in others;
 CC Name=1;
 CC IsoId=P07589-1; Sequence=Displayed;
 CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
 by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
 forms) made by fibroblasts, epithelial and other cell types, is
 deposited as fibrils in the extracellular matrix.
 CC -!- PTM: Sulfated (By similarity).
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -!- SIMILARITY: Contains 15 fibronectin type III domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K00800; AAA30521.2; -
 DR PIR; A26452; FNBO.
 DR HSP; P08253; LKS0.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibrnctnl.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN III-like.
 DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR003962; FNIII_subd.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.

DR Pfam; PF00041; fn3; 15.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00014; FNTYPEIII.
 DR ProDom; PD000995; FN Type II; 2.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.
 DR PROSITE; PS00853; FN3; 15.
 KW Acute phase; Alternative splicing; Cell adhesion; Cell shape;
 KW Glycine-protein; Heparin-binding;
 KW Direct protein sequencing; Glycoprotein; Heparin-binding;
 KW Phosphorylation; Plasma; Pyridoxone carboxylic acid; Repeat;
 KW Sulfation.
 FT MOD_RES 1 1 Pyridoxone carboxylic acid.
 FT DOMAIN 21 241 Fibrin- and heparin-binding 1.
 FT DOMAIN 277 577 Collagen-binding.
 FT DNA_BIND 876 1141 Cell-attachment.
 FT DOMAIN 1236 1509 Heparin-binding 2.
 FT DOMAIN 1600 1870 Fibrin-binding 2.
 FT DOMAIN 1991 2216 Fibronection type-I 1.
 FT DOMAIN 19 59 Fibronection type-I 2.
 FT DOMAIN 64 107 Fibronection type-I 3.
 FT DOMAIN 108 151 Fibronection type-I 4.
 FT DOMAIN 153 197 Fibronection type-I 5.
 FT DOMAIN 198 242 Fibronection type-I 6.
 FT DOMAIN 275 314 Fibronection type-I 7.
 FT DOMAIN 314 373 Fibronection type-I 8.
 FT DOMAIN 374 438 Fibronection type-I 9.
 FT DOMAIN 437 480 Fibronection type-I 10.
 FT DOMAIN 485 527 Fibronection type-I 11.
 FT DOMAIN 528 571 Fibronection type-I 12.
 FT DOMAIN 576 668 Fibronection type-III 1.
 FT DOMAIN 689 767 Fibronection type-III 2.
 FT DOMAIN 780 868 Fibronection type-III 3.
 FT DOMAIN 877 964 Fibronection type-III 4.
 FT DOMAIN 965 1053 Fibronection type-III 5.
 FT DOMAIN 1056 1141 Fibronection type-III 6.
 FT DOMAIN 1142 1234 Fibronection type-III 7.
 FT DOMAIN 1235 1325 Fibronection type-III 8.
 FT DOMAIN 1326 1415 Fibronection type-III 9.
 FT DOMAIN 1416 1505 Fibronection type-III 10.
 FT DOMAIN 1510 1599 Fibronection type-III 11 (extra domain).
 FT DOMAIN 1602 1689 Fibronection type-III 12.
 FT DOMAIN 1692 1780 Fibronection type-III 13.
 FT DOMAIN 1781 1870 Fibronection type-III 14.
 FT DOMAIN 1871 1930 Connecting strand 3 (CS-3) (V region).
 FT DOMAIN 1979 2069 Fibronection type-III 15.
 FT DOMAIN 2083 2127 Fibronection type-I 10.
 FT DOMAIN 2128 2170 Fibronection type-I 11.
 FT DOMAIN 2172 2215 Fibronection type-I 12.
 FT SITE 1493 1495 Cell attachment site.
 FT DISULFID 21 47
 FT DISULFID 45 56
 FT DISULFID 66 94
 FT DISULFID 92 104
 FT DISULFID 110 138
 FT DISULFID 136 148
 FT DISULFID 155 184
 FT DISULFID 182 194
 FT DISULFID 200 229
 FT DISULFID 227 239
 FT DISULFID 277 304
 FT DISULFID 302 311
 FT DISULFID 329 355
 FT DISULFID 343 370
 FT DISULFID 389 415
 FT DISULFID 403 430
 FT DISULFID 439 467
 FT DISULFID 465 477
 FT DISULFID 487 514
 FT DISULFID 512 524
 FT DISULFID 530 558
 FT DISULFID 556 568

```

FT DISULFID 2085 2114
FT DISULFID 2112 2124
FT DISULFID 2130 2157
FT DISULFID 2155 2167
FT DISULFID 2174 2200
FT DISULFID 2198 2209
FT DISULFID 2246 2246
FT DISULFID 2250 2250
FT MOD_RES 845 845
FT MOD_RES 850 850
FT CARBOHYD 399 399
FT CARBOHYD 497 497
FT CARBOHYD 511 511
FT CARBOHYD 846 846
FT CARBOHYD 976 976
FT CARBOHYD 1213 1213
FT CARBOHYD 1987 1987
FT CARBOHYD 1943 1943
FT CARBOHYD 1944 1944
FT MOD_RES 2263 2263
SQ SEQUENCE 2265 AA; 249557 MW; C2D21D486F498D5C CRC64;

Query Match 96.5%; Score 1447; DB 1; Length 2265;
Best Local Similarity 95.4%; Pred. No. 1.4e-116;
Matches 250; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAOQWVQSPVAVSQSPKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 1 QAOQIVQSPVAVSQSPKPGCYDNGKHQYQINQOWERTYLGVALVCTCYGSGRGFNCESKP 60

QY 61 EAEETCFDKYGTNTYRVGDTYERPKDSIMWDCICGAGRGISCTIANRCHGGQSYKIG 120
Db 61 EPEETCFDKYGTNTYRVGDTYERPKDSIMWDCICGAGRGISCTIANRCHGGQSYKIG 120

QY 121 DTWRPHETGGVLMSCVCLNGKGGWTKCPAIEKCFDHAAGTSYVVGETWKPQGWMMV 180
Db 121 DTWRPHETGGVLMSCVCLNGKGGWTKCPAIEKCFDHAAGTSYVVGETWKPQGWMMV 180

QY 181 DCTCLGEGSGRITCTSRNCNDODRTSYRIGDGTWSKDNRLQLQCTGNRGWKCE 240
Db 181 DCTCLGEGSGRITCTSRNCNDODRTSYRIGDGTWSKDNRLQLQCTGNRGWKCE 240

QY 241 RHTSIVQTTSSGSGPFTDRAAV 262
Db 241 RHTSIVQTTSSGSGPFTDRAAV 262

RESULT 11
FINC MOUSE
ID FINC MOUSE STANDARD; PRT: 2477 AA.
AC P11276; Q61567; Q61569; Q64233; Q80UI4;
DT 01-JUL-1989 (Rel. 11, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE Fibronectin precursor (FN).
GN Name=Fn1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-920 FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Breast tumor;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 1-28 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94131313; PubMed=8299972; DOI=10.1016/0378-1119(93)90036-3;
RA Polly P., Nicholson R.C.;
RT "Sequence of the mouse fibronectin-encoding gene promoter region.";
RL Gene 137:353-354 (1993).
RN [3]
RP SEQUENCE OF 562-834 FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=95403556; PubMed=7673336;
RA Tait J.F., Weller A., Timpl R., Ekblom M., Ekblom P.;
RT "Regulation of mesenchymal extracellular matrix protein synthesis by
RT transforming growth factor-beta and glucocorticoids in tumor stroma.";
RL J. Cell Sci. 108:2153-2162 (1995).
RN [4]
RP SEQUENCE OF 999-2376 FROM N.A.
RA Gorski G., Aros M., Norton P.;
RL Submitted (Dec-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 2375-2477 FROM N.A.
RX MEDLINE=88124987; PubMed=3124113;
RA Blatti S.P., Foster D.N., Ranganathan G., Moses H.L., Getz M.J.;
RT "Induction of fibronectin gene transcription and mRNA is a primary
RT response to growth-factor stimulation of AKR-2B cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1119-1123 (1988).
RN [6]
RP SEQUENCE OF 2375-2477 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93011702; PubMed=1327855;
RA Khandjian E.W., Salomon C., Leonard N., Tremblay S., Turler H.;
RT "Fibronectin gene expression in proliferating, quiescent, and SV40-
RT infected mouse kidney cells.";
RL Exp. Cell Res. 202:464-470 (1992).
RN [7]
RP STRUCTURE BY NMR OF 1447-1630.
RX MEDLINE=98202578; PubMed=9533887; DOI=10.1006/jmbi.1998.1616;
RA Copie V., Tomita Y., Akiyama S.K., Aota S., Yamada K.M., Venable R.M.,
RA Pastor R.W., Krueger S., Torchia D.A.;
RT "Solution structure and dynamics of linked cell attachment modules of
RT mouse fibronectin containing the RGD and synergy regions: comparison
RT with the human fibronectin crystal structure.";
RL J. Mol. Biol. 277:663-682 (1998).
RN [8]
RP DOWN-REGULATION BY GLUCOCORTICOIDS.
RX MEDLINE=21600963; PubMed=11737251;
RA Gu Y.-C., Tait J.F., Gullberg D., Timpl R., Ekblom M.;
RT "Glucocorticoids down-regulate the extracellular matrix proteins
RT fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma.";
RL Eur. J. Haematol. 67:176-184 (2001).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins
CC are involved in cell adhesion, cell motility, opsonization, wound
CC healing, and maintenance of cell shape.
CC -!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
CC variants, connected by 2 disulfide bonds near the carboxyl ends;
CC to a lesser extend homodimers. Interacts with FBLN1 and LGALS3BP
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced. Each of the "extra

```


CC domain" and the connecting strand 3 are present in some forms of
 CC fibronectin and absent in others;
 CC Name=1;
 CC isoId=P11276-1; Sequences-Displayed;
 CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
 CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
 CC forms) made by fibroblasts, epithelial and other cell types, is
 CC deposited as fibrils in the extracellular matrix.
 CC -!- INDUCTION: Glucocorticoids suppressed mRNA expression and protein
 CC synthesis.
 CC -!- PTM: Sulfated (By similarity).
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; BC051082; AAH51082.1; -;
 CC EMBL; Z22729; CAA80422.1; -;
 CC EMBL; X82402; CAA57796.1; -;
 CC EMBL; X93167; CAA63654.1; -;
 CC EMBL; M18194; AAA37636.1; -;
 CC EMBL; S45680; AAB23491.1; -;
 CC PIR; A49173; A49173;
 CC PIR; I48349; I48349;
 CC PDB; IMFN; NMR; @=1446-1630.
 CC PDB; 2MFN; NMR; @=1446-1630.
 CC MGD; MGI:95566; Fn1.
 CC GO; GO:0007155; P:cell adhesion; IDA.
 CC GO; GO:0007044; P:cell-substrate junction assembly; IDA.
 CC GO; GO:0042060; P:wound healing; IMP.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR000083; Fibnctnl.
 CC InterPro; IPR003962; FnIII subd.
 CC InterPro; IPR003961; FN III.
 CC InterPro; IPR008957; FN III-like.
 CC InterPro; IPR000562; FN_Type_II.
 CC Pfam; PF00039; fn1; 12.
 CC Pfam; PF00040; fn2; 2.
 CC Pfam; PF00041; fn3; 17.
 CC PRINTS; PR00014; FNTYPEIII.
 CC ProDom; PD00095; FN_Type_II; 2.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01253; FIBRONECTIN_1; 12.
 CC PROSITE; PS00023; FIBRONECTIN_2; 2.
 CC PROSITE; PS00853; FN3; 17.
 CC 3D-structure; Acute phase; Alternative splicing; Cell adhesion;
 CC Glycoprotein; Heparin-binding; Phosphorylation; Plasma; Repeat;
 CC Signal; Sulfation.
 CC SIGNAL 1 32 By similarity.
 CC CHAIN 33 2477 Fibronectin.
 CC DOMAIN 53 273 Fibrin- and heparin-binding 1.
 CC DOMAIN 308 608 Collagen-binding.
 CC DNA_BIND 906 1171
 CC DOMAIN 1357 1630 Cell-attachment.
 CC DOMAIN 1811 2081 Heparin-binding 2.
 CC DOMAIN 2296 2427 Fibrin-binding 2.
 CC DOMAIN 51 96 Fibronectin type-I 1.
 CC DOMAIN 96 140 Fibronectin type-I 2.
 CC DOMAIN 140 185 Fibronectin type-I 3.
 CC DOMAIN 185 230 Fibronectin type-I 4.
 CC DOMAIN 230 272 Fibronectin type-I 5.
 CC DOMAIN 306 343 Fibronectin type-I 6.
 CC DOMAIN 345 404 Fibronectin type-II 1.
 CC DOMAIN 405 469 Fibronectin type-II 2.
 CC DOMAIN 468 516 Fibronectin type-I 7.
 CC DOMAIN 516 559 Fibronectin type-I 8.

FT	DOMAIN	559	602	Fibronectin type-I 9.
FT	DOMAIN	607	699	Fibronectin type-III 1.
FT	DOMAIN	719	808	Fibronectin type-III 2.
FT	DOMAIN	810	897	Fibronectin type-III 3.
FT	DOMAIN	905	994	Fibronectin type-III 4.
FT	DOMAIN	995	1083	Fibronectin type-III 5.
FT	DOMAIN	1091	1171	Fibronectin type-III 6.
FT	DOMAIN	1172	1264	Fibronectin type-III 7.
FT	DOMAIN	1265	1355	Fibronectin type-III 8 (extra domain 1).
FT	DOMAIN	1356	1446	Fibronectin type-III 9.
FT	DOMAIN	1447	1536	Fibronectin type-III 10.
FT	DOMAIN	1537	1626	Fibronectin type-III 11.
FT	DOMAIN	1631	1720	Fibronectin type-III 12.
FT	DOMAIN	1721	1810	Fibronectin type-III 13 (extra domain 2).
FT	DOMAIN	1813	1900	Fibronectin type-III 14.
FT	DOMAIN	1903	1991	Fibronectin type-III 15.
FT	DOMAIN	1992	2081	Fibronectin type-III 16.
FT	DOMAIN	2082	2201	Connecting strand 3 (CS-3) (V region).
FT	DOMAIN	2190	2280	Fibronectin type-III 17.
FT	DOMAIN	2294	2338	Fibronectin type-I 10.
FT	DOMAIN	2339	2381	Fibronectin type-I 11.
FT	DOMAIN	2383	2426	Fibronectin type-I 12.
FT	SITE	1614	1616	Cell attachment site.
FT	SITE	2181	2183	By similarity.
FT	DISULFID	53	79	By similarity.
FT	DISULFID	77	88	By similarity.
FT	DISULFID	98	126	By similarity.
FT	DISULFID	124	136	By similarity.
FT	DISULFID	142	170	By similarity.
FT	DISULFID	168	180	By similarity.

Query Match 94.7%; Score 1419.5; DB 1; Length 2477;
 Best Local Similarity 93.1%; Pred. NO. 3.7e-114;
 Matches 244; Conservative 11; Mismatches 6; Indels 1; Gaps 1;

QY	1	QAOQVQPOS	PVAVSQSPGCV	DNGKHQI	NOQWERTY	LGVLVCT	CYGSGRGNCE	SKP	60
Db	33	QAOQVQPOS <td>PVAVSQSPGCV <td>DNGKHQI <td>NOQWERTY <td>LGVLVCT <td>CYGSGRGNCE <td>SKP</td> <td>92</td> </td></td></td></td></td>	PVAVSQSPGCV <td>DNGKHQI <td>NOQWERTY <td>LGVLVCT <td>CYGSGRGNCE <td>SKP</td> <td>92</td> </td></td></td></td>	DNGKHQI <td>NOQWERTY <td>LGVLVCT <td>CYGSGRGNCE <td>SKP</td> <td>92</td> </td></td></td>	NOQWERTY <td>LGVLVCT <td>CYGSGRGNCE <td>SKP</td> <td>92</td> </td></td>	LGVLVCT <td>CYGSGRGNCE <td>SKP</td> <td>92</td> </td>	CYGSGRGNCE <td>SKP</td> <td>92</td>	SKP	92
QY	61	EAEETCFDKY <td>TGNTYR <td>PKDSMI <td>WDC</td> <td>TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>120</td> </td></td></td></td>	TGNTYR <td>PKDSMI <td>WDC</td> <td>TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>120</td> </td></td></td>	PKDSMI <td>WDC</td> <td>TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>120</td> </td></td>	WDC	TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>120</td> </td>	ISCTIANR <td>CHGGQSYKIG</td> <td>120</td>	CHGGQSYKIG	120
Db	93	EPEETCFDKY <td>TGNTYR <td>PKDSMI <td>WDC</td> <td>TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>152</td> </td></td></td></td>	TGNTYR <td>PKDSMI <td>WDC</td> <td>TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>152</td> </td></td></td>	PKDSMI <td>WDC</td> <td>TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>152</td> </td></td>	WDC	TCIGAGR <td>ISCTIANR <td>CHGGQSYKIG</td> <td>152</td> </td>	ISCTIANR <td>CHGGQSYKIG</td> <td>152</td>	CHGGQSYKIG	152
QY	121	DTWRPHETGG <td>YMLCVC <td>LGNGKGE <td>WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>180</td> </td></td></td></td></td></td>	YMLCVC <td>LGNGKGE <td>WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>180</td> </td></td></td></td></td>	LGNGKGE <td>WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>180</td> </td></td></td></td>	WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>180</td> </td></td></td>	PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>180</td> </td></td>	HDHAAGTS <td>YVVGTEW <td>180</td> </td>	YVVGTEW <td>180</td>	180
Db	153	DKWRRPHETGG <td>YMLCVC <td>LGNGKGE <td>WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>212</td> </td></td></td></td></td></td>	YMLCVC <td>LGNGKGE <td>WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>212</td> </td></td></td></td></td>	LGNGKGE <td>WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>212</td> </td></td></td></td>	WTCP <td>PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>212</td> </td></td></td>	PIAEKCP <td>HDHAAGTS <td>YVVGTEW <td>212</td> </td></td>	HDHAAGTS <td>YVVGTEW <td>212</td> </td>	YVVGTEW <td>212</td>	212
QY	181	DCICLGEGSG <td>GRITCT <td>SRNRCND</td> <td>QDT</td> <td>RTSYRIG</td> <td>DTWSKDN</td> <td>RGNLLQCI</td> <td>240</td> </td>	GRITCT <td>SRNRCND</td> <td>QDT</td> <td>RTSYRIG</td> <td>DTWSKDN</td> <td>RGNLLQCI</td> <td>240</td>	SRNRCND	QDT	RTSYRIG	DTWSKDN	RGNLLQCI	240
Db	213	DCICLGEGSG <td>GRITCT <td>SRNRCND</td> <td>QDT</td> <td>RTSYRIG</td> <td>DTWSKDN</td> <td>RGNLLQCI</td> <td>272</td> </td>	GRITCT <td>SRNRCND</td> <td>QDT</td> <td>RTSYRIG</td> <td>DTWSKDN</td> <td>RGNLLQCI</td> <td>272</td>	SRNRCND	QDT	RTSYRIG	DTWSKDN	RGNLLQCI	272
QY	241	RHTSVQTTSS <td>SGSGPFTD</td> <td>VRAAV</td> <td>262</td> <td></td> <td></td> <td></td> <td></td>	SGSGPFTD	VRAAV	262				
Db	273	RH-AQSA	SAGSGS	FTDVR	TAI	293			

RESULT 12
 FINC RAT
 ID FINC RAT STANDARD; PRT; 2477 AA.
 AC P04937; Q6LXD9;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Fibronectin precursor (FN).
 GN Name=Fn1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer; TISSUE=Liver;
 RX MEDLINE=88054951; PubMed=2445560;

RA Schwarzbauer J.E., Patel R.S., Fonda D., Hynes R.O.;
RT "Multiple sites of alternative splicing of the rat fibronectin gene
transcript.";
RL EMBO J. 6:2573-2580 (1987).
RN [2]
RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=88054950; PubMed=3119323;
RA Patel R.S., Odermatt E., Schwarzbauer J.E., Hynes R.O.;
RT "Organization of the fibronectin gene provides evidence for exon
shuffling during evolution.";
RL EMBO J. 6:2565-2572 (1987).
RN [3]
RP SEQUENCE OF 1586-2477 FROM N.A.
RX MEDLINE=84082067; PubMed=6317187; DOI=10.1016/0092-8674(83)90175-7;
RA Schwarzbauer J.E., Tamkun J.W., Lemischka I.R., Hynes R.O.;
RT "Three different fibronectin mRNAs arise by alternative splicing
within the coding region.";
RL Cell 35:421-431 (1983).
RN [4]
RP SEQUENCE OF 1722-1810 FROM N.A., AND ALTERNATIVE SPLICING.
RX PubMed=3863113;
RA Odermatt E., Tamkun J.W., Hynes R.O.;
RT "Repeating modular structure of the fibronectin gene: relationship to
protein structure and subunit variation.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6571-6575 (1985).
RN [5]
RP SEQUENCE OF 2052-2237 FROM N.A., AND ALTERNATIVE SPLICING.
RX PubMed=6089177;
RA Tamkun J.W., Schwarzbauer J.E., Hynes R.O.;
RT "A single rat fibronectin gene generates three different mRNAs by
alternative splicing of a complex exon.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5140-5144 (1984).
RN [6]
RP SEQUENCE OF 1183-1192; 1385-1399 AND 2287-2300, AND INTERACTION WITH
AMPB.
RX PubMed=7519849;
RA Falkenberg C., Enghild J.J., Thøgersen I.B., Salvesen G.,
RA Akerstrom B.;
RT "Isolation and characterization of fibronectin-alpha 1-microglobulin
complex in rat plasma.";
RL Biochem. J. 301:745-751 (1994).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
including collagen, fibrin, heparin, DNA, and actin. Fibronectins
are involved in cell adhesion, cell motility, opsonization, wound
healing, and maintenance of cell shape.
CC -!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
variants, connected by 2 disulfide bonds near the carboxyl ends;
to a lesser extent homodimers. Interacts with FBLN1, AMPB and
LGALS3BP (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Each of the "extra domain" and the connecting strand 3
are present in some forms of fibronectin and absent in others;
Name=1;
CC IsoId=P04937-1; Sequence=Displayed;
CC Name=2; Synonyms=FNII-13-less;
CC IsoId=P04937-2; Sequence=VSP_003258;
CC Name=3; Synonyms=Lambda-RLF4-5;
CC IsoId=P04937-3; Sequence=VSP_003259;
CC Name=4; Synonyms=Lambda-RLF6;
CC IsoId=P04937-4; Sequence=VSP_003260;
CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
forms), made by fibroblasts, epithelial and other cell types, is
deposited as fibrils in the extracellular matrix.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
CC -!- SIMILARITY: Contains 17 fibronectin type III domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X15906; CAA34020.1; -;
DR EMBL; L29191; AAA41166.1; -;
DR EMBL; L00191; AAA41166.1; JOINED.
DR EMBL; L29191; AAA41167.1; -;
DR EMBL; L00191; AAA41167.1; JOINED.
DR EMBL; L29191; AAA41168.1; -;
DR EMBL; L00191; AAA41168.1; JOINED.
DR EMBL; M11750; AAA41170.1; -;
DR EMBL; X05831; CAA29278.1; -;
DR EMBL; X05832; CAA29279.1; -;
DR EMBL; X05833; CAA29280.1; -;
DR EMBL; X05834; CAA29281.1; -;
DR PIR; S14428; S14428.
DR HSP; P08253; 1KS0.
DR RGD; 2624; Fn1.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003962; FnIII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR PROSITE; PS00022; EGF_1_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS08053; FN3; 17.
KW Acute phase; Alternative splicing; Cell adhesion;
KW Direct protein sequencing; Glycoprotein; Heparin-binding;
KW Phosphorylation; Plasma; Repeat; Signal; Sulfation.
FT SIGNAL 1 32
FT CHAIN 33 2477 Fibronectin.
FT DOMAIN 53 273 Fibrin- and heparin-binding 1.
FT DOMAIN 308 608 Collagen-binding.
FT DNA_BIND 906 1171
FT DOMAIN 1357 1630 Cell-attachment.
FT DOMAIN 1811 2081 Heparin-binding 2.
FT DOMAIN 2296 2427 Fibrin-binding 2.
FT DOMAIN 51 91 Fibronectin type-I 1.
FT DOMAIN 96 139 Fibronectin type-I 2.
FT DOMAIN 140 183 Fibronectin type-I 3.
FT DOMAIN 185 229 Fibronectin type-I 4.
FT DOMAIN 230 274 Fibronectin type-I 5.
FT DOMAIN 306 345 Fibronectin type-I 6.
FT DOMAIN 345 404 Fibronectin type-II 1.
FT DOMAIN 405 469 Fibronectin type-II 2.
FT DOMAIN 468 511 Fibronectin type-I 7.
FT DOMAIN 516 558 Fibronectin type-I 8.
FT DOMAIN 559 602 Fibronectin type-I 9.
FT DOMAIN 607 699 Fibronectin type-III 1.
FT DOMAIN 719 808 Fibronectin type-III 2.
FT DOMAIN 810 897 Fibronectin type-III 3.
FT DOMAIN 907 994 Fibronectin type-III 4.
FT DOMAIN 995 1083 Fibronectin type-III 5.
FT DOMAIN 1091 1171 Fibronectin type-III 6.
FT DOMAIN 1172 1264 Fibronectin type-III 7.
FT DOMAIN 1265 1355 Fibronectin type-III 8 (extra domain 1).
FT DOMAIN 1356 1446 Fibronectin type-III 9.
FT DOMAIN 1447 1536 Fibronectin type-III 10.
FT DOMAIN 1537 1626 Fibronectin type-III 11.
FT DOMAIN 1631 1720 Fibronectin type-III 12.
FT DOMAIN 1721 1810 Fibronectin type-III 13 (extra domain 2).
FT DOMAIN 1721 1810

FT	DOMAIN	1813	1900	Fibronectin type-III 14.
FT	DOMAIN	1903	1991	Fibronectin type-III 15.
FT	DOMAIN	1992	2081	Fibronectin type-III 16.
FT	DOMAIN	2082	2201	Connecting strand 3 (CS-3) (V region).
FT	DOMAIN	2190	2280	Fibronectin type-III 17.
FT	DOMAIN	2294	2338	Fibronectin type-I 10.
FT	DOMAIN	2339	2381	Fibronectin type-I 11.
FT	DOMAIN	2383	2426	Fibronectin type-I 12.
FT	SITE	1614	1616	Cell attachment site.
FT	SITE	2181	2183	By similarity.
FT	DISULFID	53	79	By similarity.
FT	DISULFID	77	88	By similarity.
FT	DISULFID	98	126	By similarity.
FT	DISULFID	124	136	By similarity.
FT	DISULFID	142	170	By similarity.
FT	DISULFID	168	180	By similarity.
FT	DISULFID	187	216	By similarity.
FT	DISULFID	214	226	By similarity.
FT	DISULFID	232	261	By similarity.
FT	DISULFID	259	271	By similarity.
FT	DISULFID	308	335	By similarity.
FT	DISULFID	333	342	By similarity.
FT	DISULFID	360	386	By similarity.
FT	DISULFID	374	401	By similarity.
FT	DISULFID	420	446	By similarity.
FT	DISULFID	434	461	By similarity.
FT	DISULFID	470	498	By similarity.
FT	DISULFID	496	508	By similarity.
FT	DISULFID	518	545	By similarity.
FT	DISULFID	543	555	By similarity.
FT	DISULFID	561	589	By similarity.
FT	DISULFID	587	599	By similarity.
FT	DISULFID	2296	2325	By similarity.
FT	DISULFID	2323	2335	By similarity.
FT	DISULFID	2341	2368	By similarity.
FT	DISULFID	2366	2378	By similarity.
FT	DISULFID	2385	2409	By similarity.
FT	DISULFID	2407	2423	By similarity.
FT	DISULFID	2458	2458	Interchain (with C-2462).
Query Match				94.2%; Score 1412.5; DB 1; Length 2477;
Best Local Similarity				92.7%; Pred. No. 1.5e-113;
Matches 243; Conservative 10; Mismatches 8; Indels 1; Gaps 1;				
QY	1	QAOQMVPQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNNVLVCTCYGSGRGFNCSKP	60	
Db	33	QAOQIVQPPSPVAVSQSKPGCFDNGKHQYQINQOWERTYLGNNVLVCTCYGSGRGFNCSKP	92	
QY	61	EABETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG	120	
Db	93	EPEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG	152	
QY	121	DTWRRPHETGGYMLECVCLGNGKGWEVTCPIAEKCFDHAAGTSYVVGTEWKEPYQGMWV	180	
Db	153	DKWRRPHETGGYMLECVCLGNGKGWEVTCPIAEKCFDHAAGTSYVVGTEWKEPYQGMWV	212	
QY	181	DCTCLGSGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKDNRGNLLQCICTGNRGWEKCE	240	
Db	213	DCTCLGEGNGRITCTSRNRCNDQDTRTSYRIGDTWSKKDNRGNLLQCVCCTGNRGWEKCE	272	
QY	241	RHTSVQTTSSGSGPFTDVRVAV	262	
Db	273	RHV-LQSASAGSGSFTDVRTAI	293	
RESULT 13				
Q8C6J7	PRELIMINARY; PRT; 296 AA.			
AC	Q8C6J7			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-			

length enriched library, clone:E330027109 product:fibronectin 1, full insert sequence.

DE Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RP [1]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;

RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [4]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;

RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M., Sui M., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK054456; BAC35784.1; -.

DR HSSP; Q96KE7; 1FBR.

DR GO; GO:0005576; C:extracellular; IEA.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR000083; Fibnctnl.

DR Pfam; PF00039; fnl; 5.

DR PRINTS; PR00012; FNTPEI.

```
DR SMART; SM00058; FN1; 5.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 5.
SQ SEQUENCE 296 AA; 32617 MW; A0FF5F4809FB6439 CRC64;

Query Match 92.2%; Score 1382.5; DB 2; Length 296;
Best Local Similarity 92.0%; Pred. No. 76-112;
Matches 240; Conservative 10; Mismatches 8; Indels 3; Gaps 2;

QY 1 QAAQVQSPVAVSQSGKGYDNGKHQVINOQWERTYLVNVLCTCYGSGRGFNCESKP 60
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33 QAAQVQSPVAVSQSGKGYDNGKHQVINOQWERTYLVNVLCTCYGSGRGFNCESKP 92
QY 61 EAETCFDKYGTNTYRVDYTPERPKDSMIWDCICAGRGRICTIANRCHGGGQSYKIG 120
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 EPEETCFDKYGTNTYRVDYTPERPKDSMIWDCICAGRGRICTIANRCHGGGQSYKIG 152
QY 121 DTWRPHETGGYMLBCVCLNGKGBWTCKPIAEKCFDHAAGTSYVVGETWEKPYQGMVMV 180
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 DKWRPHETGGYMLBCVCLNGKGBWTCKPIAEKCFDHAAGTSYVVGETWEKPYQGMVMV 212
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRNLLOCTGNGRGWKCE 240
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRNLLOCTGNGRGWKCE 272
QY 241 RHTSVQTTSSGS--GPFTDVR 259
DQ ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 273 RH-ALQSASAGEALGPGTAPR 292

RESULT 14
F1NC XENLA STANDARD; PRT; 2481 AA.
AC Q91740.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fibronectin precursor.
GN Name=FNI;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92111942; PubMed=1730390;
RA Desimone D.W., Norton P.A., Hynes R.O.;
RT "Identification and characterization of alternatively spliced
RL fibronectin mRNAs expressed in early Xenopus embryos.";
RL Dev. Biol. 149:357-369(1992).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins
CC are involved in cell adhesion, cell motility, opsonization, wound
CC healing, and maintenance of cell shape (By similarity).
CC -!- SUBUNIT: Dimers or multimers of alternatively spliced variants,
CC connected by 2 disulfide bonds near the carboxyl ends (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced. Each of the "extra
CC domain" and the connecting strand 3 are present in some forms of
CC fibronectin and absent in others;
CC Name=1;
CC IsoId=Q91740-1; Sequences=Displayed;
CC -!- TISSUE SPECIFICITY: In early Xenopus embryo, cellular forms of
CC fibronectin predominate which include both extra domains. In
CC fibronectin of embryonic and adult liver the connecting strand 3
CC can be either completely excluded or included.
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77820; AAA49707.1; -.
DR HSSP; P08253; 1J7M.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrnctnl.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00013; FNYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 11.
DR PROSITE; PS00853; FN3; 17.
KW Acute phase; Alternative splicing; Cell adhesion; Glycoprotein;
KW Heparin-binding; Plasma; Repeat; Signal.
FT SIGNAL 1 31
FT CHAIN 32 2481
FT DOMAIN 55 275 Fibrin- and heparin-binding 1.
FT DOMAIN 309 609 Collagen-binding.
FT DNA_BIND 907 1172 By similarity.
FT DOMAIN 1358 1631 Cell-attachment.
FT DOMAIN 1812 2082 Heparin-binding 2.
FT DOMAIN 2301 2432 Fibrin-binding 2.
FT DOMAIN 53 93 Fibronectin type-I 1.
FT DOMAIN 98 141 Fibronectin type-I 2.
FT DOMAIN 142 185 Fibronectin type-I 3.
FT DOMAIN 187 231 Fibronectin type-I 4.
FT DOMAIN 232 276 Fibronectin type-I 5.
FT DOMAIN 307 346 Fibronectin type-I 6.
FT DOMAIN 346 405 Fibronectin type-II 1.
FT DOMAIN 406 470 Fibronectin type-II 2.
FT DOMAIN 469 512 Fibronectin type-I 7.
FT DOMAIN 517 559 Fibronectin type-I 8.
FT DOMAIN 560 603 Fibronectin type-I 9.
FT DOMAIN 608 700 Fibronectin type-III 1.
FT DOMAIN 719 809 Fibronectin type-III 2.
FT DOMAIN 811 898 Fibronectin type-III 3.
FT DOMAIN 908 995 Fibronectin type-III 4.
FT DOMAIN 996 1084 Fibronectin type-III 5.
FT DOMAIN 1087 1172 Fibronectin type-III 6.
FT DOMAIN 1173 1265 Fibronectin type-III 7.
FT DOMAIN 1266 1356 Fibronectin type-III 8 (extra domain).
FT DOMAIN 1357 1447 Fibronectin type-III 9.
FT DOMAIN 1448 1537 Fibronectin type-III 10.
FT DOMAIN 1538 1627 Fibronectin type-III 11.
FT DOMAIN 1632 1721 Fibronectin type-III 12.
FT DOMAIN 1722 1811 Fibronectin type-III 13 (extra domain).
FT DOMAIN 1814 1901 Fibronectin type-III 14.
FT DOMAIN 1904 1992 Fibronectin type-III 15.
FT DOMAIN 1993 2082 Fibronectin type-III 16.
FT DOMAIN 2083 2205 Connecting strand 3 (CS-3) (V region).
FT DOMAIN 2206 2284 Fibronectin type-III 17.
FT DOMAIN 2292 2343 Fibronectin type-I 10.
FT DOMAIN 2344 2386 Fibronectin type-I 11.
FT DOMAIN 2388 2431 Fibronectin type-I 12.
FT SITE 1615 1617 Cell attachment site.
FT DISULFID 55 81 By similarity.
FT DISULFID 79 90 By similarity.
FT DISULFID 100 128 By similarity.
FT DISULFID 126 138 By similarity.
```

```
FT DISULFID 144 172 By similarity.
FT DISULFID 170 182 By similarity.
FT DISULFID 189 218 By similarity.
FT DISULFID 216 228 By similarity.
FT DISULFID 234 263 By similarity.
FT DISULFID 261 273 By similarity.
FT DISULFID 309 336 By similarity.
FT DISULFID 334 343 By similarity.
FT DISULFID 361 387 By similarity.
FT DISULFID 375 402 By similarity.
FT DISULFID 421 447 By similarity.
FT DISULFID 471 499 By similarity.
FT DISULFID 487 509 By similarity.
FT DISULFID 519 546 By similarity.
FT DISULFID 544 556 By similarity.
FT DISULFID 562 590 By similarity.
FT DISULFID 598 600 By similarity.
FT DISULFID 2301 2330 By similarity.
FT DISULFID 2328 2340 By similarity.
FT DISULFID 2346 2373 By similarity.
FT DISULFID 2371 2383 By similarity.
FT DISULFID 2390 2414 By similarity.
FT DISULFID 2412 2428 By similarity.
FT DISULFID 2459 2459 Interchain (with C-2463) (By similarity).
FT DISULFID 2463 2463 Interchain (with C-2459) (By similarity).
FT CARBOHYD 431 431 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 529 529 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 543 539 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 877 877 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1244 1244 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1291 1291 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2202 2202 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 2481 AA; 272678 MW; 7E47DF4F6CE72C93 CRC64;

Query Match 86.5%; Score 1296; DB 1; Length 2481;
Best Local Similarity 83.2%; Pred. No. 1.9e-103;
Matches 218; Conservative 22; Mismatches 20; Indels 2; Gaps 1;

QY 1 QAOQVVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGTVLCTCYGSGRGFNCESKP 60
Db 37 QQQQVVQPG--TQDNHQKGYDNGKGYQINQWERTYLGTVLCTCYGSGRGFNCESKP 94

QY 61 EAEETCFDKYTGNTVRVGTYPKDSMIWDCITCAGGRISCTIANRCHGGGQSYKIG 120
Db 95 ESEETCFDKYTGVSRYVGETYRPKDNMIWDCITCAGGRISCTIANRCHGGGQSYKIG 154

QY 121 DTRRPHEGTGGMLECVCLGNGKGEWTKPIAEKFDHAGTSYVVGWETWKPQGMVMV 180
Db 155 DTRRPHEGTGGMLECVCLGNGKGEWTKPVAERCYDNTAGTSYVVGWETWKPQGMVMV 214

QY 181 DCTCLGEGSGRITCSRNRCNDQDTRTSYRIGDTWSKDNRNGLLQICITGNRGWEKCE 240
Db 215 DCTCLGEGNGRITCSSKNCNDQDTRTSYRIGDTWSKDTNRGNLLQICITGNRGWEKCE 274

QY 241 RHTSVQTTSSGSGPFTDRAAV 262
Db 275 RHSSAQTGSGNPITNIOTAL 296

RESULT 15
Q6GQA5 PRELIMINARY; PRT; 2481 AA.
AC Q6GQA5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fibronectin protein.
GN Name=fibronectin;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Fetingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko M., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Griwood J., Schmitz J., Myers R.M., Butterfield V.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schmerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391(2002).

[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072841; AAH72841.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 11.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00853; FN3; 17.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
SQ SEQUENCE 2481 AA; 272708 MW; 4DB47E5656B8B5A3 CRC64;
```

```
Query Match 86.5%; Score 1296; DB 2; Length 2481;
Best Local Similarity 83.2%; Pred. No. 1.9e-103;
Matches 218; Conservative 22; Mismatches 20; Indels 2; Gaps 1;

QY 1 QAOQVVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGTVLCTCYGSGRGFNCESKP 60
Db 37 QQQQVVQPG--TQDNHQKGYDNGKGYQINQWERTYLGTVLCTCYGSGRGFNCESKP 94

QY 61 EAEETCFDKYTGNTVRVGTYPKDSMIWDCITCAGGRISCTIANRCHGGGQSYKIG 120
Db 95 ESEETCFDKYTGVSRYVGETYRPKDNMIWDCITCAGGRISCTIANRCHGGGQSYKIG 154
```

QY	121	DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV	180
Db	155	DTWRRPHETGGYMLECVCLGNGKGWTCCKPVAERCYDNTAGTSYVVGQIWEKPYQGMMV	214
QY	181	DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKONRGNLLQCICTGNRGWEKCE	240
Db	215	DCTCLGEGNGRITCSSKNRCNDQDTRTSYRIGDTWSKTDTRGNLLQCICTGNRGWEKCE	274
QY	241	RHTSVQTTSSGSGPFTDVRAAV	262
Db	275	RHSSAQATGTGNSNPITNIQTAL	296

Search completed: September 30, 2005, 13:37:55
Job time : 191.55 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 13:30:51 : Search time 108.675 Seconds
(without alignments)
932.422 Million cell updates/sec

Title: US-09-492-971B-15_COPY_4_265
Perfect score: 1499
Sequence: 1 QAAQMVQPSPVAVSQSKPG.....TSVQTTSGSGPFTDRAAV 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499	100.0	720	2 AAY28914	Aay28914 Fibronect
2	1499	100.0	2324	2 AAR92778	Aar92778 Human fib
3	1499	100.0	2324	5 AAU74674	Aau74674 Human fib
4	1499	100.0	2324	5 AAE23651	Aae23651 Human pro
5	1499	100.0	2328	4 AAG68182	Aag68182 Fibronect
6	1499	100.0	2328	6 ABU07486	Abu07486 Protein d
7	1499	100.0	2328	6 ABR41106	Abra41106 Human fib
8	1499	100.0	2328	6 ABR92078	Abra92078 Human cer
9	1499	100.0	2328	7 ADB70378	Adb70378 Fibronect
10	1499	100.0	2328	7 ADB98726	Adb98726 Human fib
11	1499	100.0	2328	7 ADB82522	Adb82522 Human pro
12	1499	100.0	2328	8 ADJ37157	Adj37157 Human mal
13	1499	100.0	2386	2 AAW63171	Aaw63171 Amino aci
14	1499	100.0	2446	2 AAR60021	Aar60021 Fibrinoge
15	1499	100.0	2477	2 AAW99595	Aaw99595 Human fib
16	1496	99.8	1179	8 ADP75952	Adp75952 Human min
17	1496	99.8	1359	8 ADP75957	Adp75957 Human leu
18	1495	99.7	463	6 ABR58303	Abra58303 BCU0707 p
19	1495	99.7	642	8 ADQ39403	Adq39403 Human myo
20	1495	99.7	642	8 ADQ67316	Adq67316 Human bla
21	1495	99.7	642	8 ADS17489	Ads17489 Amino aci
22	1495	99.7	642	8 ADR97658	Adr97658 Human fib
23	1495	99.7	657	8 ADQ39409	Adq39409 Human myo
24	1495	99.7	660	2 AAY28901	Aay28901 Human mig
25	1495	99.7	984	8 ADQ39406	Adq39406 Human myo

26	1495	99.7	2182	8 ADR66462	Adr66462 Human pro
27	1495	99.7	2182	8 ADR66120	Adr66120 Human pro
28	1495	99.7	2220	6 ABO01289	Ab001289 Human NOV
29	1495	99.7	2220	4 ADN95950	Adn95950 Human NOV
30	1495	99.7	2265	4 AAM38647	Aam38647 Human pol
31	1495	99.7	2266	6 ABR40124	Abra40124 Human cel
32	1495	99.7	2296	8 ADO39412	Ado39412 Human myo
33	1495	99.7	2330	4 AAM38646	Aam38646 Human pol
34	1495	99.7	2355	4 AAM38649	Aam38649 Human pol
35	1495	99.7	2355	6 ABR58335	Abra58335 NM_00202
36	1495	99.7	2355	7 ADP65196	Adp65196 Human fib
37	1495	99.7	2355	8 ADG89560	Adg89560 Human fib
38	1495	99.7	2355	8 ADL92160	Adl92160 Fibronect
39	1495	99.7	2355	8 ADO55175	Ado55175 Protein #
40	1495	99.7	2355	8 ADQ26085	Adq26085 Fibronect
41	1495	99.7	2355	8 ADQ29668	Adq29668 Human col
42	1495	99.7	2355	8 ADQ39415	Adq39415 Human myo
43	1495	99.7	2355	8 ADQ39404	Adq39404 Human myo
44	1495	99.7	2355	8 ADR67315	Adr67315 Human bla
45	1495	99.7	2386	4 AAM38648	Aam38648 Human pol

ALIGNMENTS

RESULT 1
AAY28914
ID AAY28914 standard; protein; 720 AA.

XX
AC AAY28914;
XX
DT 21-SEP-1999 (first entry)
XX
DE Fibronectin protein sequence.
XX
KW Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epicope; fibronectin.
XX
OS Homo sapiens.
XX
PN WO9931233-Al.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-GB003766.
XX
PR 16-DEC-1997; 97GB-00026539.
XX
(UYDU-) UNIV DUNDEE.
PI Schor SL, Schor AM;
XX
WPI; 1999-430039/36.
XX
PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.
XX
PS Disclosure; Fig 2; 86pp; English.
XX
CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence CC represents the human fibronectin

SQ Sequence 720 AA;

Query Match 100.0%; Score 1499; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.9e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQMVQPSPVAVSQSKPGCYDNGKHVQINQWERTYILGNVLVCTCYGSGRGNCSKP 60

```

Db 50 QAOQWVQSPVAVSQSGPCYDNGKHQIQQWERTYLGNVLVCTCYGSRGFNCESKP 109
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGGRISCTIANRCHGGQSYKIG 120
Db 110 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGGRISCTIANRCHGGQSYKIG 169
QY 121 DTWRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEPYQGWMMV 180
Db 170 DTWRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEPYQGWMMV 229
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQCICTGNRGWEKCE 240
Db 230 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQCICTGNRGWEKCE 289
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 290 RHTSVQTTSSGSGPFTDVRAAV 311

```

RESULT 2

```

AAR92778
ID AAR92778 standard; protein; 2324 AA.
XX AC
XX 21-OCT-2004 (revised)
DT 21-JUN-1996 (first entry)
XX
DE Human fibronectin.
XX
KW Fibrin-binding peptide; fibronectin; fibrinolysis; fibrinogenesis;
KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis; therapy.
XX
OS Homo sapiens.
XX Unidentified.
XX
FH Key
FT Domain
FT 21..65
FT /label = 1F1
FT /note= "type 1 module 1"
FT 66..109
FT /label = 2F1
FT /note= "type 1 module 2"
FT 110..154
FT /label = 3F1
FT /note= "type 1 module 3"
FT 155..199
FT /label = 4F1
FT /note= "type 1 module 4"
FT 200..244
FT /label = 5F1
FT /note= "type 1 module 5"
FT 277..312
FT /label = 6F1
FT /note= "type 1 module 6"
FT 439..486
FT /label = 7F1
FT /note= "type 1 module 7"
FT 439..467
FT /note= "Cys439-Cys467 disulfide bond"
FT 465..477
FT /note= "Cys465-Cys477 disulfide bond"
FT 487..529
FT /label = 8F1
FT /note= "type 1 module 8"
FT 530..569
FT /label = 9F1
FT /note= "type 1 module 9"
FT 2123..2230
FT /label = Fibrin_binding_domain
FT /note= "11 kDa C-terminal fibrin-binding domain"
FT 2141..2185

```

```

FT /label = 10F1
FT /note= "type 1 module 10"
FT Disulfide-bond 2144..2173
FT /note= "Cys2144-Cys2173 disulfide bond"
FT Disulfide-bond 2171..2183
FT /note= "Cys2171-Cys2183 disulfide bond"
FT Domain 2187..2230
FT /label = 11F1
FT /note= "type 1 module 11"
FT Disulfide-bond 2189..2216
FT /note= "Cys2189-Cys2216 disulfide bond"
FT Disulfide-bond 2214..2226
FT /note= "Cys2214-Cys2226 disulfide bond"
FT Domain 2233..2271
FT /label = 12F1
FT /note= "type 1 module 12"
XX
XX WO9604304-A1.
XX
XX 15-FEB-1996.
XX
XX 01-AUG-1995; 95WO-US009819.
XX
XX 01-AUG-1994; 94US-00283857.
XX (UJNY ) UNIV NEW YORK STATE.
XX
XX Gold LI, Rostagno AA;
PI
XX
XX WPI; 1996-129333/13.
XX N-PSDB; AAT17551.
XX
XX New fibrin-binding peptide molecules - used for the diagnosis and
XX treatment of conditions associated with fibrin deposition, e.g. thrombi.
XX
XX Claim 2; Page 104-110; 146pp; English.
XX
XX Human fibronectin (AAR92778) has a fibrin-binding site, close to the C-
XX terminus (amino acids 2123-2232 or 2141-2230), covering the 10F1.11F1
XX module pair of the protein. Fibrin-binding sites can be prep. by
XX enzymatic cleavage of fibronectin, peptide synthesis or by recombinant
XX DNA techniques. They are used to detect a fibrin-binding target site, to
XX treat disorders involving abnormal fibrinolysis or fibrinogenesis, to
XX disrupt blood clots and to aid delivery of medicaments to fibrin- contg.
XX sites. They show high affinity to and slow dissociation from fibrin, and
XX provide fast diffusion and rapid clearance
XX
XX Revised record issued on 21-OCT-2004 : Correction to Feature Table Key
XX
XX Sequence 2324 AA;
XX
Query Match 100.0%; Score 1499; DB 2; Length 2324;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAOQWVQSPVAVSQSGPCYDNGKHQIQQWERTYLGNVLVCTCYGSRGFNCESKP 60
Db 1 QAOQWVQSPVAVSQSGPCYDNGKHQIQQWERTYLGNVLVCTCYGSRGFNCESKP 60
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGGRISCTIANRCHGGQSYKIG 120
Db 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGGRISCTIANRCHGGQSYKIG 120
QY 121 DTWRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEPYQGWMMV 180
Db 121 DTWRPHETGGYMLCEVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEPYQGWMMV 180
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQCICTGNRGWEKCE 240
Db 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQCICTGNRGWEKCE 240
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262

```



```
Db      241 RHTSVQTTSSGSGPFTDVRAAV 262
RESULT 3
AAU74674
ID      AAU74674 standard; protein; 2324 AA.
XX
AC      AAU74674;
XX
DT      09-APR-2002 (first entry)
XX
DE      Human fibronectin protein.
XX
KW      Human; fibronectin; VLA-4 binding site; very late antigen-4;
KW      quiescent haematopoietic cell; HC; apoptosis; CD34+.
XX
OS      Homo sapiens.
XX
PN      WO200187071-A1.
XX
PD      22-NOV-2001.
XX
PF      12-MAY-2000; 2000WO-US012993.
XX
PR      12-MAY-2000; 2000WO-US012993.
XX
PA      (ADRE-) ADVANCED RES & TECHNOLOGY INST.
XX
PI      Williams DA, Bradford GB, Dutt P, Yoder MC;
XX
DR      WPI; 2002-082932/11.
XX
PT      Obtaining hematopoietic cell population containing quiescent cells for
PT      use in treating a subject, by expanding the cells while adhered to a
PT      polypeptide containing binding sites for integrins on the cells.
XX
PS      Disclosure; Page 51-61; 63pp; English.
XX
CC      The invention relates to obtaining a population of quiescent
CC      haematopoietic cells (HC), comprising culturing HC while adhering the
CC      cells to a polypeptide having a very late antigen-4 (VLA-4) binding site,
CC      in particular fibronectin polypeptide, so as to expand the number of HC.
CC      where the adhering provides an increased percentage of quiescent HC. Also
CC      included are inducing apoptosis of a subpopulation of HC, by contacting
CC      the cells with a polypeptide having a VLA-4 binding site under conditions
CC      to cause apoptosis of a subpopulation of HC and a medium for culturing HC
CC      which enriches quiescent HC, comprising a fibronectin polypeptide. The
CC      method is useful for obtaining a cell population containing quiescent
CC      haematopoietic cells which are useful for treating a subject. Use of the
CC      VLA-4 containing peptide is useful for inducing apoptosis of a
CC      subpopulation of CD34+ HC population. The present sequence is human
CC      fibronectin which contains then VLA-4 binding site used in the method of
CC      the invention
XX
SQ      Sequence 2324 AA;
Query Match      100.0%; Score 1499; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db      1 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
QY      61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTTIANRCHGGQSYKIG 120
Db      61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTTIANRCHGGQSYKIG 120
QY      121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVVGTEWKFYQGMVMV 180
Db      121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVVGTEWKFYQGMVMV 180
QY      181 DCTCLGEGSGRITCTSRNRCNDQDTSYRIGDTWSKNRGNLLQICITGNRGWEKCE 240
```

```
Db      181 DCTCLGEGSGRITCTSRNRCNDQDTSYRIGDTWSKNRGNLLQICITGNRGWEKCE 240
QY      241 RHTSVQTTSSGSGPFTDVRAAV 262
Db      241 RHTSVQTTSSGSGPFTDVRAAV 262
RESULT 4
AAE23651
ID      AAE23651 standard; protein; 2324 AA.
XX
AC      AAE23651;
XX
DT      10-SEP-2002 (first entry)
XX
DE      Human protein related to modulation of proliferation of EPCs.
XX
KW      Human; proliferation; apoptosis; erythroid progenitor cell; EPC; VLA;
KW      integrin very late antigen; haematopoietic cell; induced disorder;
KW      genetic disorder; congenital erythroid disorder; polycythaemia; anaemia.
XX
OS      Homo sapiens.
XX
PN      WO200187037-A1.
XX
PD      22-NOV-2001.
XX
PF      12-MAY-2000; 2000WO-US012961.
XX
PR      12-MAY-2000; 2000WO-US012961.
XX
PA      (ADRE-) ADVANCED RES & TECHNOLOGY INST.
XX
PI      Williams DA, Kapur R, Cooper R, Zhang L;
XX
DR      WPI; 2002-121963/16.
XX
PT      Modulating proliferation of erythroid progenitor cells in a hematopoietic
PT      cellular population by culturing with polypeptides containing an integrin
PT      very late antigen 5 or 4 binding site, useful in treating erythroid
PT      disorders.
XX
PS      Disclosure; Page 55-64; 66pp; English.
XX
CC      The present invention relates to a method for enhancing proliferation or
CC      inducing apoptosis of erythroid progenitor cells (EPCs). The method
CC      involves culturing haematopoietic cellular populations containing EPCs in
CC      the presence of a polypeptide having an integrin very late antigen (VLA) -
CC      5 or VLA-4 binding site. The EPCs are useful for treating patients having
CC      induced, genetic or congenital erythroid disorders. Methods of the
CC      invention are useful in diagnostic assays of progenitor cells, e.g. cells
CC      occurring in pathologic states such as anaemia or polycythaemia and in
CC      the study of erythroid cell development. They are used in the screening
CC      of agents which inhibit or prevent apoptosis of erythroid cells. The
CC      present sequence is a human protein related to the modulation of
CC      proliferation of EPCs. This sequence is used in the invention
XX
SQ      Sequence 2324 AA;
Query Match      100.0%; Score 1499; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db      1 QAAQWVQPSVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
QY      61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTTIANRCHGGQSYKIG 120
Db      61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTTIANRCHGGQSYKIG 120
QY      121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVVGTEWKFYQGMVMV 180
```

```
Db 121 DTWRPHETGGYMLECVCLNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEXPYQGMWV 180
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWKKDNKGNLLQCICTGNRGWKE 240
Db 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWKKDNKGNLLQCICTGNRGWKE 240
QY 241 RHTSVQTTSSGSGPFTDVAHV 262
Db 241 RHTSVQTTSSGSGPFTDVAHV 262

RESULT 5
AAG68182
ID AAG68182 standard; protein; 2328 AA.
XX
AC AAG68182;
XX
DT 25-JAN-2002 (first entry)
XX
DE Fibronectin protein SEQ ID NO:98.
XX
KW Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
KW sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
KW antisense therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
KW osteomalacia; fibrous dysplasia.
XX
OS Homo sapiens.
XX
PN WO200177327-A1.
XX
PD 18-OCT-2001.
XX
PF 21-JUN-2000; 2000WO-US016951.
XX
PR 05-APR-2000; 2000US-00543771.
PR 05-APR-2000; 2000US-00544398.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Carulli JP, Little RD, Recker RR, Johnson ML;
XX
DR WPI; 2001-657171/75.
XX
PT New high bone mass (HBM) and Zmax1 genes and proteins useful for
PT modulating bone mass for the treatment of e.g. osteoporosis.
XX
PS Claim 76; Page 408-413; 443pp; English.
XX
CC The present invention describes the human Zmax1 gene and the high bone
CC mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
CC genes have osteopathic activities. The genes can be used in gene therapy,
CC antisense therapy and in the production of vaccines. They can be used in
CC the diagnosis and treatment of bone disorders including osteoporosis,
CC Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
CC ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 2328 AA;

Query Match 100.0%; Score 1499; DB 4; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QACQMVQPSVAVSQSGCGYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCEKSP 60
Db 5 QACQMVQPSVAVSQSGCGYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCEKSP 64
QY 61 EAEETCFDXYTGNTVRVGTYPKDSMTWDCCTCIGAGRGRTSCTIANRCHGGOSYKIG 120
Db 65 EAEETCFDXYTGNTVRVGTYPKDSMTWDCCTCIGAGRGRTSCTIANRCHGGOSYKIG 124
QY 121 DTWRPHETGGYMLECVCLNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEXPYQGMWV 180
```

```
Db 125 DTWRPHETGGYMLECVCLNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEXPYQGMWV 184
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWKKDNKGNLLQCICTGNRGWKE 240
Db 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWKKDNKGNLLQCICTGNRGWKE 244
QY 241 RHTSVQTTSSGSGPFTDVAHV 262
Db 245 RHTSVQTTSSGSGPFTDVAHV 266

RESULT 6
ABU07486
ID ABU07486 standard; protein; 2328 AA.
XX
AC ABU07486;
XX
DT 28-JAN-2003 (first entry)
XX
DE Protein differentially regulated in prostate cancer #89.
XX
KW Prostate cancer; gene expression; differential regulation;
KW molecular marker; drug target; cancer detection; cancer diagnosis;
KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX
OS Homo sapiens.
XX
PN WO200281638-A2.
XX
PD 17-OCT-2002.
XX
PF 08-APR-2002; 2002WO-US010824.
XX
PR 06-APR-2001; 2001US-0281731P.
PR 06-APR-2001; 2001US-0281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
DR WPI; 2003-058520/05.
DR N-PSDB; ABX10391.
XX
PT Novel genes which are differentially regulated in prostate cancer, useful
PT for diagnosing prostate cancer in prostate tissue sample and assessing
PT therapeutic or preventive intervention in prostate cancer patients.
XX
PS Claim 1; Page 378-386; 416pp; English.
XX
CC The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity. (I)
CC is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
```

CC Of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is useful
CC in therapeutic applications to treat prostate cancer. The identification
CC of specific genes, and groups of genes, expressed in pathways
CC physiologically relevant to prostate cancer permits the definition of
CC functional and disease pathways and the delineation of targets in these
CC pathways which are useful in diagnostic, therapeutic, and clinical
CC applications. This is the amino acid sequence of a protein differentially
CC regulated in prostate cancer
XX
XX Sequence 2328 AA;

Query Match 100.0%; Score 1499; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAOQWVQPOS PVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
DB 5 QAOQWVQPOS PVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 64
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 124
QY 121 DTWRRPHETGGYMLCVCVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWV 180
DB 125 DTWRRPHETGGYMLCVCVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWV 184
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNKRNLLQICITGNRGEWKCE 240
DB 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNKRNLLQICITGNRGEWKCE 244
QY 241 RHTSVQTTSSGSGPFTDVRVAV 262
DB 245 RHTSVQTTSSGSGPFTDVRVAV 266

RESULT 7
ABR41106
ID ABR41106 standard; protein; 2328 AA.
AC ABR41106;
XX
XX
XX 02-JUN-2003 (first entry)
XX Human fibronectin gene FN protein product.
XX
XX Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone.
XX
XX Homo sapiens.
XX
XX WO200292764-A2.
XX
XX 21-NOV-2002.
XX
XX 13-MAY-2002; 2002WO-US014876.
XX
XX 11-MAY-2001; 2001US-0290071P.
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP) WYETH.
XX
XX Babij P, Bex FJ, Yaworsky PJ, Bodine PV;
XX
XX WPI; 2003-129278/12.
XX

XX New transgenic animals (e.g. mice), useful as models for studying bone
PT density modulation, developing drugs for treating or preventing bone
PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
PT reduced bone density.
XX
XX Disclosure; Page 532-538; 603pp; English.
XX
XX The invention relates to novel transgenic animals expressing the high
CC bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
CC comprising an alteration of the gene encoding LRP5 or LRP6, or expressing
CC an LRP5 that is modulated by an altered gene control sequence introduced
CC by homologous or non-homologous recombination. The transgenic animals are
CC for the study of bone density modulation or bone mass modulation. The
CC invention has osteopathic and cytostatic activity. The polynucleotides of
CC the invention may have a use in gene therapy. The transgenic animals and
CC nucleic acids are for the study of bone density modulation, where the
CC bone mass is modulated relative to non-transgenic animals of the same
CC species in more than one parameter selected from bone density, bone
CC strength, trabecular number, bone size, or bone tissue connectivity. The
CC transgenic animals, nucleic acids and methods are useful for identifying
CC molecules involved in bone development, and for developing pharmaceutical
CC compositions, which may be employed for treating or preventing bone
CC diseases, e.g. osteoporosis, osteomalacia, rickets, Paget's disease, or
CC neoplasms of the bone. The transgenic animals and nucleic acids are also
CC useful in methods for diagnosing diseases involved in bone development, is
CC or characterised by reduced bone density or mass. The present sequence, is
CC used in the exemplification of the invention
XX
XX Sequence 2328 AA;

Query Match 100.0%; Score 1499; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAOQWVQPOS PVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
DB 5 QAOQWVQPOS PVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 64
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 124
QY 121 DTWRRPHETGGYMLCVCVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWV 180
DB 125 DTWRRPHETGGYMLCVCVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWV 184
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNKRNLLQICITGNRGEWKCE 240
DB 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNKRNLLQICITGNRGEWKCE 244
QY 241 RHTSVQTTSSGSGPFTDVRVAV 262
DB 245 RHTSVQTTSSGSGPFTDVRVAV 266

RESULT 8
ABR92078
ID ABR92078 standard; protein; 2328 AA.
XX
XX ABR92078;
XX
XX 10-SEP-2003 (first entry)
XX Human cervical cancer cell marker protein SEQ ID NO:64.
XX Human; cervical cancer; cervical cancer marker; cancer therapy;
KW detection; gene therapy; vaccine.
XX
XX Homo sapiens.
XX
XX WO2002101075-A2.
XX

PD 19-DEC-2002.
XX
XX
PF 12-JUN-2002; 2002WO-US018638.
XX
PR 13-JUN-2001; 2001US-0298155P.
PR 13-JUN-2001; 2001US-0298159P.
PR 14-NOV-2001; 2001US-0335936P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gamavarapu M, Glatt K, Hoersch S;
PI
XX
DR WPI; 2003-156967/15.
DR N-PSDB; ACF12859.
XX
XX New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
XX Claim 4; Page 212-217; 386pp; English.
PS
XX
XX ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer.
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterizing,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials
XX
XX Sequence 2328 AA;
SQ
Query Match 100.0%; Score 1499; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQVQPQSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 60
DB 5 QAAQVQPQSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 64
QY 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTIANRCHGGQSYKIG 124
QY 121 DTWRPHEPHTGGYMLCEVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWEPYQGMWV 180
DB 125 DTWRPHEPHTGGYMLCEVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWEPYQGMWV 184
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKDNRNLLQICITGNRGGEWKE 240
DB 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKDNRNLLQICITGNRGGEWKE 244
QY 241 RHTSVQTTSSSGPFTDVRVAV 262
DB 245 RHTSVQTTSSSGPFTDVRVAV 266
RESULT 9
ID ADB70378
XX ADB70378 standard; protein; 2328 AA.
XX
AC ADB70378;
XX
DT 04-DEC-2003 (first entry)
XX

DE Fibronectin SEQ ID NO:70.
XX
XX cancer; malignant pleural mesothelioma; MPM; lung adenocarcinoma;
KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
KW human.
XX
XX Homo sapiens.
XX
XX WO2003021229-A2.
XX
XX PD 13-MAR-2003.
XX
XX 05-SEP-2002; 2002WO-US028203.
XX
XX 05-SEP-2001; 2001US-0317389P.
PR 30-AUG-2002; 2002US-00236031.
XX
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Gordon GJ, Jensen RV, Gullans SR, Bueno R;
PI
XX
XX WPI; 2003-290233/28.
DR N-PSDB; ADB70377.
XX
XX Diagnosing cancer cells in tissue sample, or determining prognosis or
PT outcome of cancer patient, by calculating ratio of expression levels of
PT genes that are differentially expressed in cancer and non cancer tissues.
XX
XX Claim 77; Page 369-376; 396pp; English.
XX
XX The present invention describes a method (M1) for diagnosing the presence
CC of cancer cells or non-cancer cells in a tissue sample, or determining
CC the prognosis or outcome of a cancer patient. M1 involves providing a set
CC of genes that are differentially expressed in cancerous or non-cancerous
CC conditions, determining the expression levels of the set of genes and
CC calculating a ratio of the expression levels of the differentially
CC expressed genes. M1 is useful for diagnosing the presence of cancer cells
CC or non-cancer cells in a tissue sample, where the cancer is malignant
CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma,
CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell
CC lymphoma, follicular lymphoma and ovarian cancer, and for determining
CC prognosis or outcome of a cancer patient. The ratio of expression levels
CC of differentially expressed genes is used as an indicator of cancer type,
CC cancer class, and/or cancer prognosis, all of which are useful for
CC determining a course of treatment of a patient. The present sequence
CC represents a human protein which is used in an example from the present
CC invention.
XX
XX Sequence 2328 AA;
SQ
Query Match 100.0%; Score 1499; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQVQPQSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 60
DB 5 QAAQVQPQSPVAVSQSGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 64
QY 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTIANRCHGGQSYKIG 124
QY 121 DTWRPHEPHTGGYMLCEVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWEPYQGMWV 180
DB 125 DTWRPHEPHTGGYMLCEVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWEPYQGMWV 184
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKDNRNLLQICITGNRGGEWKE 240
DB 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDWTWSKDNRNLLQICITGNRGGEWKE 244
QY 241 RHTSVQTTSSSGPFTDVRVAV 262
|||||

Db 245 RHTSVQTTSSGSGPFTDRAAV 266

RESULT 10
ADB98726
ID ADB98726 standard; protein; 2328 AA.
XX
AC ADB98726;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human fibronectin.
XX
KW Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;
bone mass modulation; osteoporosis.
XX
OS Homo sapiens.
XX
PN WO200292000-A2.
XX
PD 21-NOV-2002.
XX
PF 13-MAY-2002; 2002WO-US014877.
XX
PR 11-MAY-2001; 2001US-0290071P.
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP) WYETH.
XX
PI Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;
XX
DR WPI; 2003-129214/12.
DR P-PSDB; ADB98703.
XX
PT New nucleic acid comprising a mutation in LRP5 or LRP6, useful for
diagnosing a HBM-like phenotype in a subject and for preparing a
composition for modulating bone mass and/or lipid levels in a subject
suffering from e.g. osteoporosis.
XX
PS Disclosure; SEQ ID NO 677; 629pp; English.
XX
CC The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and
LRP6 mutants, which results in a HBM-like phenotype when expressed in a
cell. The HBM-like phenotype results in bone mass modulation and/or lipid
level modulation. The invention is useful for diagnosing a HBM-like
phenotype in a subject and for preparing a composition for modulating
bone mass and/or lipid levels in a subject suffering from e.g.
osteoporosis. The present sequence was used to illustrate the invention.
XX
SQ Sequence 2328 AA;

Query Match 100.0%; Score 1499; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQQMVQPSFVAVSQSKPCYDNGKHQYQINQWERTYLGNIIVCTCYGSGRGFNCSKP 60
DB 5 QAQQMVQPSFVAVSQSKPCYDNGKHQYQINQWERTYLGNIIVCTCYGSGRGFNCSKP 64

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 124

QY 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 180
DB 125 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 184

QY 181 DCTCLGEGSGRICTSRNRNCDDTTRTSYRIGTWSKDNRLLOCICTGNRGRGWKCE 240
DB 185 DCTCLGEGSGRICTSRNRNCDDTTRTSYRIGTWSKDNRLLOCICTGNRGRGWKCE 244

QY 241 RHTSVQTTSSGSGPFTDRAAV 262
DB 245 RHTSVQTTSSGSGPFTDRAAV 266

RESULT 11
ADE82522
ID ADE82522 standard; protein; 2328 AA.
XX
AC ADE82522;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein sequence related to the invention #12.
XX
KW LRP5; LRP6; HBM; Dkk activity; Osteopathic; Antiinflammatory;
Antiarthritic; bone mass disorders; osteoporosis; hypercalcaemia;
hyperostosis; osteogenesis; Wnt signaling.
XX
OS Homo sapiens.
XX
PN WO200292015-A2.
XX
PD 21-NOV-2002.
XX
PF 17-MAY-2002; 2002WO-US015982.
XX
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP) WYETH.
XX
PI Allen K, Anisowicz A, Bhat BM, Damagnez V, Robinson JA;
PI Yaworsky PJ;
XX
DR WPI; 2003-129219/12.
XX
PT Regulating LRP5, LRP6 or HBM activity in a subject, useful for modulating
lipid levels and/or bone mass, and for in treating bone mass disorders,
e.g. osteoporosis, comprises administering a composition which modulates
a Dkk activity.
XX
PS Disclosure; SEQ ID NO 98; 173pp; English.
XX
CC The present invention relates to regulating LRP5, LRP6 or HBM activity in
a subject comprising administering a composition which modulates a Dkk
activity. The method is useful for modulating lipid levels and/or bone
mass, and is useful in treating or diagnosing abnormal lipid levels and
bone mass disorders, such as osteoporosis, bone fracture, age-related
loss of bone, a chondrodystrophy, drug-induced bone disorder, high bone
turnover, hypercalcaemia, hyperostosis, osteogenesis, imperfecta,
osteomalacia, osteomyelitis, Paget's disease, osteoarthritis, and
rickets. Modulators of Dkk activity are useful for as reagents in
studying bone mass and lipid level modulation, in modulating Wnt
signaling, or treating Dkk-mediated disorders. The present sequence
represents a human protein sequence related to the invention.
XX
SQ Sequence 2328 AA;

Query Match 100.0%; Score 1499; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.8e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQQMVQPSFVAVSQSKPCYDNGKHQYQINQWERTYLGNIIVCTCYGSGRGFNCSKP 60
DB 5 QAQQMVQPSFVAVSQSKPCYDNGKHQYQINQWERTYLGNIIVCTCYGSGRGFNCSKP 64

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 124

XX Sequence 2386 AA;
SQ
Query Match 100.0%; Score 1499; DB 2; Length 2386;
Best Local Similarity 100.0%; Pred. No. 7e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAOQWVQPOS PAVASQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
DB 32 QAOQWVQPOS PAVASQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 91
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHEGGQSYKIG 120
DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHEGGQSYKIG 151
QY 121 DTWRRPHETGGYMLCEVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 180
DB 152 DTWRRPHETGGYMLCEVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 211
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRLQICITGNRGGEWKCE 240
DB 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRLQICITGNRGGEWKCE 271
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 272 RHTSVQTTSSGSGPFTDVRAAV 293
RESULT 14
AAR60021
ID AAR60021 standard; protein; 2446 AA.
XX AC AAR60021;
XX
DT 25-MAR-2003 (revised)
DT 23-FEB-1995 (first entry)
XX
DE Fibrinogen-alpha.
XX
KW Tissue binding; tissue sealing; wound healing; vulnery;
KW tissue-binding domain; TSD; crosslinking domain; fibronectin;
KW fibrinogen-alpha; heparin-binding domain; collagen-binding domain;
KW cell-binding domain; hybrid protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 282..608
FT /label= collagen-binding domain
FT /note= "acts as tissue-binding domain of hybrid protein"
FT 1812..2171
FT /label= heparin-binding domain
FT /note= "acts as tissue-binding domain of hybrid protein"
XX
PN WO9416085-A2.
XX
XX
PD 21-JUL-1994.
XX
XX 30-DEC-1993; 93WO-US012687.
XX
XX 30-DEC-1992; 92US-00998271.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Irani MH;
XX
XX WPI; 1994-249231/30.
DR N-PSDB; AAQ70009.
XX
XX New hybrid proteins for use in tissue sealing and wound healing -
PT comprising a tissue-binding domain from a protein covalently linked to a
PT crosslinking domain of another protein.
XX

PS Disclosure; Page 37-48; 87pp; English.
XX
CC Hybrid proteins have a tissue-binding domain (TBD) from 1 protein linked
to a crosslinking domain from another protein. The TBD comprises: aa 2-
926, 928-1338 and especially 2-1336 of the sequence given in AAR60019;
the heparin-binding domain (aa 1812-2171 of AAR60021) of fibronectin; the
collagen-binding domain (aa 282-608 of AAR60021) of fibronectin; or the
cell-binding domain (aa 1357-1903 or 1532- 1631 of AAR60020) of
fibronectin. DNA encoding a fibronectin- fibrinogen hybrid is given in
CC AAQ70007, and sequences for fibronectin and fibrinogen-alpha in AAQ70008
and AAQ70009, respectively. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2446 AA;
Query Match 100.0%; Score 1499; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 7.2e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAOQWVQPOS PAVASQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
DB 32 QAOQWVQPOS PAVASQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 91
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHEGGQSYKIG 120
DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHEGGQSYKIG 151
QY 121 DTWRRPHETGGYMLCEVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 180
DB 152 DTWRRPHETGGYMLCEVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 211
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRLQICITGNRGGEWKCE 240
DB 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRLQICITGNRGGEWKCE 271
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 272 RHTSVQTTSSGSGPFTDVRAAV 293
RESULT 15
AAW99595
ID AAW99595 standard; peptide; 2477 AA.
XX AC AAW99595;
XX
DT 22-JUN-1999 (first entry)
XX
DE Human fibronectin.
XX
XX Human; fibronectin; glycoprotein; extracellular matrix; motif; migration;
stimulation; wound healing; periodontal tissue regeneration; metastasis;
tumour; invasion; angiogenesis; inflammation; connective tissue function.
XX
OS Homo sapiens.
XX
XX WO9902674-A1.
XX
XX 21-JAN-1999.
XX
XX 01-JUL-1998; 98WO-GB001939.
XX
XX 08-JUL-1997; 97GB-00014276.
XX
XX (UYDU-) UNIV DUNDEE.
XX
XX Schor SL;
XX
XX WPI; 1999-120875/10.
XX
XX New peptides which modulate cell migration have the motif IGD - and are
PT useful to stimulate angiogenesis.
XX
XX Disclosure; Fig 5; 57pp; English.
PS

XX This sequence represents the amino acid sequence of human fibronectin, a
CC widely distributed glycoprotein present in extracellular matrices. The
CC protein contains the motifs IGS (AAW99593) and IGDQ (AAW99594) which can
CC be used in peptides to stimulate cell migration for wound healing,
CC periodontal tissue regeneration, angiogenesis, inhibition of tumour
CC invasion and metastasis, and in relation to inflammation or connective
CC tissue function
XX
SQ Sequence 2477 AA;

Query Match 100.0%; Score 1499; DB 2; Length 2477;
Best Local Similarity 100.0%; Pred. No. 7.3e-109;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QAOQMVQPSQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCEKSP 60
Db 32 QAOQMVQPSQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCEKSP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTIANRCHEGGOSYKIG 120
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTIANRCHEGGOSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
Db 152 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 211
Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKDNRGNLLQICITGNRGWEKCE 240
Db 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKDNRGNLLQICITGNRGWEKCE 271
Qy 241 RHTSVQTTSSSGSPFTDVRAAV 262
Db 272 RHTSVQTTSSSGSPFTDVRAAV 293

Search completed: September 30, 2005, 13:34:05
Job time : 110.675 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 30, 2005, 13:30:51 ; Search time 40.4373 Seconds
(without alignments)
623.404 Million cell updates/sec

Title: US-09-492-971b-15_copy_4_265
Perfect score: 1499
Sequence: 1 QAAQMVQSPVAVSQSKPG.....TSVQTTSSGSGPFTDVRRAV 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499	100.0	2386	1 FNHU	fibronectin precur
2	1447	96.5	2265	1 FNBO	fibronectin - bovi
3	1412.5	94.2	2477	2 S14428	fibronectin - precu
4	1296	86.5	2481	2 A43908	fibronectin - Afri
5	269	17.9	190	2 I51279	fibronectin - east
6	228.5	15.2	1020	2 A29355	fibronectin - chic
7	164	10.9	103	2 A49173	fibronectin - mous
8	138	9.2	1220	2 A56136	jagged protein pre
9	135	9.0	5376	2 T42215	zonadhesin - mouse
10	128.5	8.6	473	2 A56175	adhesive plaque pr
11	121	8.1	2139	2 A35672	crumbs protein - f
12	118.5	7.9	1810	1 A32230	tenascin precursor
13	116.5	7.8	1584	2 T22674	hypothetical prote
14	115	7.7	1808	2 T15099	hypothetical prote
15	114	7.6	1790	1 MRPB1	laminin beta-1 cha
16	113.5	7.6	1557	2 T28811	hypothetical prote
17	112.5	7.5	2331	2 S18188	notch protein homo
18	112.5	7.5	2555	2 A40043	Notch homolog prot
19	112	7.5	2352	2 T30201	Xotch protein - Af
20	111	7.4	2524	2 A35844	cell-fate determin
21	110	7.3	2471	2 A49128	EGF repeat transme
22	109.5	7.3	1897	2 T30176	hypothetical prote
23	109	7.3	550	2 T47158	transmembrane prot
24	109	7.3	2437	2 S42612	zonadhesin - pig
25	108	7.2	2476	2 T34028	protein T22A3.8 [i
26	107.5	7.2	2823	2 F87908	hypothetical prote
27	107.5	7.2	2823	2 T23064	hypothetical prote
28	107.5	7.2	3102	2 T43291	laminin alpha chain
29	107.5	7.2	3106	1 S53868	laminin alpha-2 ch

30	107	7.1	1700	2 S08167	Balbani ring 3 pr
31	105	7.0	1964	2 T09059	notch4 - mouse
32	105	7.0	2531	2 T31070	notch homolog - se
33	104.5	7.0	955	2 A45441	thrombospondin 4 -
34	104.5	7.0	1722	2 E89753	protein Fl1c7.4 [i
35	104	6.9	306	2 T26639	hypothetical prote
36	104	6.9	1372	2 T25933	hypothetical prote
37	104	6.9	2201	2 A32160	tenascin-C - human
38	104	6.9	2531	2 A46019	notch-1 protein -
39	104	6.9	3084	1 MMMSA	laminin alpha-1 ch
40	103.5	6.9	485	2 S36772	E-selectin - bovin
41	103.5	6.9	768	2 B41029	integrin beta-8 ch
42	103.5	6.9	2871	2 A55624	fibrillin-1 precu
43	103.5	6.9	3002	2 A47221	fibrillin-1 precu
44	103	6.9	577	2 B37057	integrin beta-6 ch
45	103	6.9	3707	2 S18252	heparan sulfate pr

ALIGNMENTS

RESULT 1

FNHU
fibronectin precursor [validated] - human
N:Alternate names: fibronectin splice form ED-A
C:Species: Homo sapiens (man)
C>Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A26284
R:Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A:Reference number: A26460; MUID:87175578; PMID:3031656
A:Accession: A26460
A:Molecule type: DNA
A:Residues: 1-49 <DEA>
A:Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA533
R:Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A:Title: Evolution of the fibronectin gene.
A:Reference number: A26284; MUID:86111901; PMID:3003095
A:Accession: A26284
A:Molecule type: DNA
A:Residues: 1447-1540 <OLD>
A:Cross-references: GB:M12549; NID:g182688
A:Note: The authors translated the codon TTC for residue 1494 as Glu
R:Paolletta, G.; Henschliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-
A:Reference number: S00848; MUID:88233940; PMID:3375063
A:Accession: S03917
A:Molecule type: DNA
A:Residues: 1594-1767, 'V', 1769-1783 <PAO>
A:Cross-references: EMBL:X07718; NID:g31402
A:Note: the authors translated the codon AAC for residue 1631 as Asp
R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene
A:Reference number: A24854; MUID:87030929; PMID:3770201
A:Accession: A24854
A:Molecule type: DNA
A:Residues: 1992-2147 <VIB>
A:Cross-references: GB:X04530; NID:g31436
R:Gutman, A.; Yamada, K.M.; Kornblitt, A.
FEBS Lett. 207, 145-148, 1986
A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
A:Reference number: A24476; MUID:87030890; PMID:3770189
A:Accession: A24476
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-14, 'Q', 16-38 <GUT>
R:Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A:Title: Primary structure of human fibronectin: differential splicing may generate at

A;Reference number: A91008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080,2112-2386 <OR>
A;Cross-references: GB:X0761
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080,2112-2386 <O2>
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:X0055; NID:gl82680; PIDN:AAAS2459.1; PID:gl82683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386 <BER>
A;Cross-references: GB:M10905; NID:gl82696; PIDN:AAAS2462.1; PID:gl82697
R;Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UME>
A;Cross-references: GB:M27589; NID:gl82705; PIDN:AAAS2465.1; PID:gl82706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991,2017-2039 <UM2>
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different me
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SEK>
A;Cross-references: GB:M14060; NID:gl82701; PIDN:AAAS2464.1; PID:gl82704
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <O3>
A;Cross-references: GB:X00799; NID:gl82681; PIDN:AAAS2460.1; PID:gl82684
R;Garcia-Pardo, A.; Pearlstain, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:840322463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300;551-560 <GAR2>

R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'O',679-703,'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2081,2113-2244 <GAR3>
R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the p
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstain, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080,2112-2356 <GAR4>
R;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
action, and transformation.
C;Genetics:
A;Gene: GDB:FN1
A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-q34
A;Introns: 49/3, 1266/1, 1357/1, 1487/1, 1541/1, 1631/1, 1721/1, 1991/1, 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F;1-26/Domain: signal sequence #status predicted <Sig>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <PHB>
F;52-87/Domain: fibronectin type I repeat homology <IF1>
F;97-135/Domain: fibronectin type I repeat homology <IF2>
F;141-179/Domain: fibronectin type I repeat homology <IF3>
F;186-225/Domain: fibronectin type I repeat homology <IF4>
F;231-270/Domain: fibronectin type I repeat homology <IF5>
F;308-608/Domain: collagen binding <CB>
F;308-342/Domain: fibronectin type I repeat homology <IF6>
F;360-401/Domain: fibronectin type II repeat homology <2FI>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <IF7>
F;518-555/Domain: fibronectin type I repeat homology <IF8>
F;561-599/Domain: fibronectin type I repeat homology <IF9>
F;609-692/Domain: fibronectin type III repeat homology <3FA>
F;616-706/Domain: heparin binding <HPB>
F;719-801/Domain: fibronectin type III repeat homology <3FB>
F;810-891/Domain: fibronectin type III repeat homology <3FC>

```
F:906-988/Domain: fibronectin type III repeat homology <3FD>
F:996-1077/Domain: fibronectin type III repeat homology <3FE>
F:1086-1164/Domain: fibronectin type III repeat homology <3FF>
F:1173-1258/Domain: fibronectin type III repeat homology <3FG>
F:1266-1349/Domain: fibronectin type III repeat homology <3FH>

Query Match      100.0%; Score 1499; DB 1; Length 2386;
Best Local Similarity 100.0%; Pred. No. 3.4e-105;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAQMVQPSVAVSQSPKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 60
Db 32 QAAQMVQPSVAVSQSPKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 91
QY 61 EAETCFDXYTGNTRYVGDYTERPKDSMTWDCCTCAGAGRGISCTTANRCHGGQSYKIG 120
Db 92 EAETCFDXYTGNTRYVGDYTERPKDSMTWDCCTCAGAGRGISCTTANRCHGGQSYKIG 151
QY 121 DTWRSPHETGGYMLKCVCLGNKGKGEWTKPIAEKCFDHAAGTSYVVGTEWXPYQGMWV 180
Db 152 DTWRSPHETGGYMLKCVCLGNKGKGEWTKPIAEKCFDHAAGTSYVVGTEWXPYQGMWV 211
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKDNRGNLLQICITGNRGFWKCE 240
Db 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKDNRGNLLQICITGNRGFWKCE 271
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 2
FNBO
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A26452; B21165; A23292
R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A:Title: Complete primary structure of bovine plasma fibronectin.
A:Reference number: A26452; MUID:87054047; PMID:3780752
A:Accession: A26452
A:Molecule type: protein
A:Residues: 1-2265 <SKO>
A:Cross-references: UNIPROT:P07589
R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; MUID:83221567; PMID:6304699
A:Accession: B21165
A:Molecule type: mRNA
A:Residues: 2170-2265 <KOR>
A:Cross-references: GB:K00800; NID:G163055; PIDN:AAA30521.2; PID:G5713323
R:Petersen, T.E.; Thogersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottrup
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal
A:Reference number: A23292; MUID:83117805; PMID:6218503
A:Accession: A23292
A:Molecule type: protein
A:Residues: 1-16,'C',18-20,'S',22-432;447-463;1367-1517;1567-1673;2062-2176,'N',2178-226
A:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.
C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b
C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
aling, and maintenance of cell shape.
C:Comment: Plasma fibronectin is synthesized by hepatocytes.
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu
F:21-241/Domain: fibrin and heparin binding <FBR>
F:21-56/Domain: fibronectin type I repeat homology <1F1>
F:66-104/Domain: fibronectin type I repeat homology <1F2>
F:110-148/Domain: fibronectin type I repeat homology <1F3>
F:155-194/Domain: fibronectin type I repeat homology <1F4>
F:200-239/Domain: fibronectin type I repeat homology <1F5>
```

```
F:277-577/Domain: collagen binding <CBR>
F:277-311/Domain: fibronectin type I repeat homology <1F6>
F:329-370/Domain: fibronectin type II repeat homology <2F1>
F:389-430/Domain: fibronectin type II repeat homology <2F2>
F:439-477/Domain: fibronectin type I repeat homology <1F7>
F:487-524/Domain: fibronectin type I repeat homology <1F8>
F:530-568/Domain: fibronectin type I repeat homology <1F9>
F:578-661/Domain: fibronectin type III repeat homology <FN3A>
F:688-770/Domain: fibronectin type III repeat homology <FN3B>
F:779-860/Domain: fibronectin type III repeat homology <FN3C>
F:875-957/Domain: fibronectin type III repeat homology <FN3D>
F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F:1233-1318/Domain: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F:1410-1517/Domain: cell attachment <CAD>
F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F:1493-1495/Region: cell attachment (R-G-D) motif
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F:1600-1870/Domain: heparin binding <HB2>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3L>
F:1781-1863/Domain: fibronectin type III repeat homology <FN3M>
F:1970-1972/Region: cell attachment (R-G-D) motif
F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F:1985-2216/Domain: fibrin binding <FB2>
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experim
F:2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted
F:399,497,511,846,976,1213,1987/Binding site: carboxylate (Asn) (covalent) #status exp
F:1205,1692/Binding site: carboxylate (Asn) (covalent) #status absent
F:1943,1944/Binding site: carboxylate (Thr) (covalent) #status experimental
F:2246/Disulfide bonds: interchain (to 2250) #status predicted
F:2250/Disulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match      96.5%; Score 1447; DB 1; Length 2265;
Best Local Similarity 95.4%; Pred. No. 3.8e-105;
Matches 250; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAAQMVQPSVAVSQSPKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSEKP 60
Db 1 QAAQIVQPSPLTVSQSPKPCYDNGKHQYQINQWERTYLGSLVCTCYGSGRGFNCSEKP 60
QY 61 EAETCFDXYTGNTRYVGDYTERPKDSMTWDCCTCAGAGRGISCTTANRCHGGQSYKIG 120
Db 61 EPEETCFDXYTGNTRYVGDYTERPKDSMTWDCCTCAGAGRGISCTTANRCHGGQSYKIG 120
QY 121 DTWRSPHETGGYMLKCVCLGNKGKGEWTKPIAEKCFDHAAGTSYVVGTEWXPYQGMWV 180
Db 121 DTWRSPHETGGYMLKCVCLGNKGKGEWTKPIAEKCFDQAAGTSYVVGTEWXPYQGMWV 180
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKDNRGNLLQICITGNRGFWKCE 240
Db 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKDNRGNLLQICITGNRGFWKCE 240
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 241 RHTSLQTTSGSGSFDTVRAAI 262

RESULT 3
S14428
fibronectin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R:Hynes, R.O.
```

submitted to the EMBL Data Library, July 1989

A:Reference number: S14428
 A:Accession: S14428
 A:Molecule type: mRNA
 A:Residues: 1-2477 <HYN>
 A:Cross-references: UNIPROT:P04937; EMBL:X15906; NID:g56164
 R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
 EMBO J. 6, 2573-2580, 1987
 A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
 A:Reference number: S12455; MUID:88054951; PMID:2445560
 A:Accession: S12455
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 609-1810, 'T', 1812-2283 <SCH>
 A:Cross-references: EMBL:X15906
 R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
 A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing
 A:Reference number: A22319; MUID:84298097; PMID:6089177
 A:Accession: A22319
 A:Molecule type: DNA
 A:Residues: 2052-2237 <TAM>
 R:Falkenberg, C.; Enghild, J.J.; Thogersen, I.B.; Salvesen, G.; Akerstroem, B.
 Biochem. J. 301, 745-751, 1994
 A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in rat liver
 A:Reference number: S46203; MUID:94330948; PMID:7519849
 A:Accession: S46203
 A:Status: Preliminary
 A:Molecule type: protein
 A:Residues: 1183-1192; 'GLN', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PY', 1385-1399 <PAL>
 R:Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.
 EMBO J. 6, 2565-2572, 1987
 A:Title: Organization of the fibronectin gene provides evidence for exon shuffling during evolution
 A:Reference number: S00459; MUID:88054950; PMID:3119323
 A:Accession: S00459
 A:Molecule type: DNA
 A:Residues: 1-139; 2382-2477 <PAT>
 A:Cross-references: EMBL:X05831
 A:Note: the authors translated the codon CCT for residues 51 and 94 as Ala
 R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.
 Cell 35, 421-431, 1983
 A:Title: Three different fibronectin mRNAs arise by alternative splicing within the coding region of the fibronectin gene
 A:Reference number: A27252; MUID:84082067; PMID:6317187
 A:Accession: A27252
 A:Molecule type: mRNA
 A:Residues: 1586-1720, 'T', 1722, 1813-2477 <SC2>
 R:Odermatt, E.; Tamkun, J.W.; Hynes, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
 A:Title: Repeating modular structure of the fibronectin gene: Relationship to protein structure
 A:Reference number: I59049; MUID:86016741; PMID:3863113
 A:Accession: I59049
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1722-1810 <RES>
 A:Cross-references: GB:M11750; NID:g204164; PIDN:AAA41170.1; PID:g554437
 C:Genetics:
 A:Introns: 51/1; 94/1; 2416/3; 2454/3
 C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology; fibronectin type I repeat homology; collagen binding; disulfide bond; duplication
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:33-2477/Product: fibronectin #status predicted <MAT>
 F:53-88/Domain: fibronectin type I repeat homology <1F1>
 F:98-136/Domain: fibronectin type I repeat homology <1F2>
 F:142-180/Domain: fibronectin type I repeat homology <1F3>
 F:187-226/Domain: fibronectin type I repeat homology <1F4>
 F:232-271/Domain: fibronectin type I repeat homology <1F5>
 F:308-342/Domain: fibronectin type I repeat homology <1F6>
 F:360-401/Domain: fibronectin type II repeat homology <2F1>
 F:420-461/Domain: fibronectin type II repeat homology <2F2>
 F:470-508/Domain: fibronectin type I repeat homology <1F7>
 F:518-555/Domain: fibronectin type I repeat homology <1F8>
 F:561-599/Domain: fibronectin type I repeat homology <1F9>
 F:609-692/Domain: fibronectin type III repeat homology <FN3A>

F:718-800/Domain: fibronectin type III repeat homology <FN3B>
 F:809-987/Domain: fibronectin type III repeat homology <FN3C>
 F:905-987/Domain: fibronectin type III repeat homology <FN3D>
 F:995-1076/Domain: fibronectin type III repeat homology <FN3E>
 F:1085-1164/Domain: fibronectin type III repeat homology <FN3F>
 F:1172-1257/Domain: fibronectin type III repeat homology <FN3G>
 F:1265-1348/Domain: fibronectin type III repeat homology <FN3H>
 F:1356-1439/Domain: fibronectin type III repeat homology <FN3I>
 F:1447-1529/Domain: fibronectin type III repeat homology <FN3J>
 F:1537-1619/Domain: fibronectin type III repeat homology <FN3K>
 F:1614-1616/Region: cell attachment (R-G-D) motif
 F:1631-1713/Domain: fibronectin type III repeat homology <FN3L>
 F:1721-1803/Domain: fibronectin type III repeat homology <FN3M>
 F:1811-1893/Domain: fibronectin type III repeat homology <FN3N>
 F:1903-1984/Domain: fibronectin type III repeat homology <FN3O>
 F:1992-2074/Domain: fibronectin type III repeat homology <FN3P>
 F:2181-2183/Region: cell attachment (R-G-D) motif
 F:2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
 F:2296-2335/Domain: fibronectin type I repeat homology <1F10>
 F:2341-2378/Domain: fibronectin type I repeat homology <1F11>
 F:2385-2420/Domain: fibronectin type I repeat homology <1F12>
 F:53-79, 77-98, 98-126, 124-136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333-368, 2365-2378, 2385-2411, 2409-2420/Disulfide bonds: #status predicted
 F:2458/Disulfide bonds: interchain (to 2462) #status predicted
 F:2462/Disulfide bonds: interchain (to 2458) #status predicted

Query Match 94.2%; Score 1412.5; DB 2; Length 2477;
 Best Local Similarity 92.7%; Pred. No. 2e-102;
 Matches 243; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

QY 1 QAAQWVQSPVAVSQKPGCYDNGKHQYINQWERTYLGNVLCVTCYGGSGFNCESKP 60
 DB 33 QAAQIVPPSPVAVSQKPGCFDNGKHQYINQWERTYLGNAVLCVTCYGGSGFNCESKP 92
 QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTIANRCHGGGOSYKIG 120
 DB 93 EPEETCFDKYTGNTYKVGDTYERPKDSMIWDCTCIGAGRISCTIANRCHGGGOSYKIG 152
 QY 121 DTWRPHETGCMLECVCLGNGKEWTCPTAEKCFDHAAGTSYVVGTEWPKYOGWMWV 180
 DB 153 DKWRPHETGTMLECLCLGNGKEWTCPTAEKCFDHAAGTSYVVGTEWPKYOGWMWV 212
 QY 181 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTMSKDKNRGLLQCITCGNGRGEWKE 240
 DB 213 DCTCLGEGSGRITCTSRNRCDQDTRTSYRIGDWTMSKDKNRGLLQCITCGNGRGEWKE 272
 QY 241 RHTSVQTTSSGSGPPTDVRVAV 262
 DB 273 RHV-LQSASAGSGSPTDVRTAI 293

RESULT 4
 A43908
 fibronectin - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
 C:Accession: A43908
 R:Desimone, D.W.; Norton, P.A.; Hynes, R.O.
 Dev. Biol. 149, 357-369, 1992
 A:Title: Identification and characterization of alternatively spliced fibronectin mRNAs
 A:Reference number: A43908; MUID:92111942; PMID:1730390
 A:Accession: A43908
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-2481 <DES>
 A:Cross-references: GB:M77820
 A:Note: sequence extracted from NCBI backbone (NCBI:77473)
 C:Superfamily: fibronectin; fibronectin type I repeat homology; glycoprotein; heterodimer
 C:Keywords: duplication; extracellular matrix; glycoprotein; heterodimer
 F:55-90/Domain: fibronectin type I repeat homology <1F1>
 F:100-138/Domain: fibronectin type I repeat homology <1F2>
 F:144-182/Domain: fibronectin type I repeat homology <1F3>
 F:189-228/Domain: fibronectin type I repeat homology <1F4>

F:234-273/Domain: fibronectin type I repeat homology <1F5>
F:309-343/Domain: fibronectin type I repeat homology <1F6>
F:361-402/Domain: fibronectin type II repeat homology <2F1>
F:421-462/Domain: fibronectin type II repeat homology <2F2>
F:471-509/Domain: fibronectin type I repeat homology <1F7>
F:519-556/Domain: fibronectin type I repeat homology <1F8>
F:562-600/Domain: fibronectin type I repeat homology <1F9>
F:610-693/Domain: fibronectin type III repeat homology <FN3A>
F:719-801/Domain: fibronectin type III repeat homology <FN3B>
F:810-891/Domain: fibronectin type III repeat homology <FN3C>
F:906-988/Domain: fibronectin type III repeat homology <FN3D>
F:996-1077/Domain: fibronectin type III repeat homology <FN3E>
F:1086-1165/Domain: fibronectin type III repeat homology <FN3F>
F:1173-1358/Domain: fibronectin type III repeat homology <FN3G>
F:1266-1349/Domain: fibronectin type III repeat homology <FN3H>
F:1357-1440/Domain: fibronectin type III repeat homology <FN3I>
F:1448-1530/Domain: fibronectin type III repeat homology <FN3J>
F:1538-1620/Domain: fibronectin type III repeat homology <FN3K>
F:1615-1617/Region: cell attachment (R-G-D) motif
F:1632-1714/Domain: fibronectin type III repeat homology <FN3L>
F:1722-1804/Domain: fibronectin type III repeat homology <FN3M>
F:1812-1894/Domain: fibronectin type III repeat homology <FN3N>
F:1904-1985/Domain: fibronectin type III repeat homology <FN3O>
F:1993-2075/Domain: fibronectin type III repeat homology <FN3P>
F:2197-2277/Domain: fibronectin type III repeat homology <FN3Q>
F:2301-2340/Domain: fibronectin type I repeat homology <1F10>
F:2346-2383/Domain: fibronectin type I repeat homology <1F11>
F:2390-2425/Domain: fibronectin type I repeat homology <1F12>
F:55-81, 79-90, 100-128, 126-138, 144-172, 170-182, 189-218, 216-228, 234-263, 261-273, 309-336, 333-343/Domain: fibronectin type I repeat homology <1F13>
F:2373-2371-2383, 2390-2416, 2414-2425/Disulfide bonds: #status predicted
F:2463/Disulfide bonds: interchain (to 2463) #status predicted
F:2463/Disulfide bonds: interchain (to 2459) #status predicted

Query Match 86.5%; Score 1296; DB 2; Length 2481;
Best Local Similarity 83.2%; Pred. No. 2, 6e-93;
Matches 218; Conservative 22; Mismatches 20; Indels 1; Gaps 1;

QY 1 QAAQWVQPSVAVSOSKPGCYDNGKHYYQINQWERTYLGNNLVCTCYGSGRGNFNCESKP 60
DB 37 QQQQVVQPG--TQDNHQKGCYDNGKYYQINQWERTYLGNTLVCTCYGSGRGNFNCESKP 94

QY 61 EAEETCFDKYGTNTYRVGTYRYPKDSMTWDCITCAGGRISCTIANRCHGGQSYKIG 120
DB 95 ESEETCFDKYGTYSRVGTYRYPKDNMTWDCITCAGGRISCTIANRCHGGQSYKIG 154

QY 121 DTWRPHETGGYMLECVCLNGKGEWTKPIAEKCFDHAAGTSYVVGWETWXPYQGWMMV 180
DB 155 DTWRPHETGGYMLECVCLNGKGEWTKPVAERCYDNTAGTSYVVGWETWXPYQGWMMV 214

QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNKGNLLQICITGNRGGEWKCE 240
DB 215 DCTCLGEGSGRITCTSSKNRCNDQDTRTSYRIGDTWSKDTNRGNLLQICITGNRGGEWKCE 274

QY 241 RHYSVQTTSSGSGPFTDVAAV 262
DB 275 RHSSAQATGSGNPITNIQTAL 296

RESULT 5
I51279
fibronectin - eastern newt (fragment)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: I51279
R:Nace, J.D.; Tassava, R.A.
Dev. Dyn. 202, 153-164, 1995
A:Title: Examination of fibronectin distribution and its sources in the regenerating newt
A:Reference number: I51279; MUID:95252528; PMID:7734733
A:Accession: I51279
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-190 <NAC>
A:Cross-references: UNIPROT:Q91400; GB:S76886; NID:G914305; PIDN:AAB34250.1; PID:G914308

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
F:7-48/Domain: fibronectin type II repeat homology <2F2>
F:57-95/Domain: fibronectin type I repeat homology <1F7>
F:105-142/Domain: fibronectin type I repeat homology <1F8>
F:148-186/Domain: fibronectin type I repeat homology <1F9>

Query Match 17.9%; Score 269; DB 2; Length 190;
Best Local Similarity 33.5%; Pred. No. 4, 3e-14;
Matches 63; Conservative 33; Mismatches 74; Indels 18; Gaps 8;

QY 65 TCFDKYGTNTYRVGDTYRYPKDSMTWDCITCIG--AGRGRISCTIANRCH-----GGQ 115
DB 6 SCHFFPLNNRYNTDCTSEGRDSMKWCGTTLNLYADQKFGCPMA--AAHEICTTNEG 63

QY 116 SYKIGDTWRPHETGGYMLECVCLNGKGEWTC---KPIAEKCFDHAAGTSYVVGWETW 172
DB 64 MYRVGVQVDKQHDQ-GHMMRCTCVNGRGEMWCVAYSQLRDQCI--VDGITVDVNHFTK 120

QY 173 PYQGMWVDCITCLGRSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNKGNLLQICICTGN 232
DB 121 RHEEGHMMNCTCYGQGRGWWKCDADQCDQDSETQFYQIGDSW-EKHVGQVRYQCYSIG 179

QY 233 GRGEWKCE 240
DB 180 GIGEWHCQ 187

RESULT 6
A29355
fibronectin - chicken (fragment)
C:Species: Gallus gallus (Chicken)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A29355
R:Norton, P.A.; Hynes, R.O.
Mol. Cell. Biol. 7, 4297-4307, 1987
A:Title: Alternative splicing of chicken fibronectin in embryos and in normal and trans
A:Reference number: A29355; MUID:88142820; PMID:2830487
A:Molecule type: mRNA
A:Residues: 1-1020 <NOR>
A:Cross-references: UNIPROT:P11722
C:Genetics:
A:Introns: 176/3

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C:Keywords: alternative splicing; duplication; extracellular matrix; glycoprotein; hete
F:1-86/Domain: fibronectin type III repeat homology <FN3H>
F:92-177/Domain: fibronectin type III repeat homology <FN3I>
F:180-262/Domain: fibronectin type III repeat homology <FN3J>
F:257-259/Region: cell attachment (R-G-D) motif
F:274-356/Domain: fibronectin type III repeat homology <FN3K>
F:364-446/Domain: fibronectin type III repeat homology <FN3L>
F:454-536/Domain: fibronectin type III repeat homology <FN3M>
F:546-628/Domain: fibronectin type III repeat homology <FN3N>
F:636-718/Domain: fibronectin type III repeat homology <FN3O>
F:837-917/Domain: fibronectin type III repeat homology <FN3P>
F:940-979/Domain: fibronectin type I repeat homology <1F10>
F:940-969, 967-979, 985-1012/Disulfide bonds: #status predicted

Query Match 15.2%; Score 228.5; DB 2; Length 1020;
Best Local Similarity 40.0%; Pred. No. 3e-10;
Matches 44; Conservative 20; Mismatches 31; Indels 15; Gaps 5;

QY 39 LGNNLVCTCYGSGRGNFNCESKPEABETCFDKYGTNTYRVGDTYRYPKDS--MIWDCITCIG 96
DB 924 VGNTV-----SEGLN----QPADDTCYDITYTGSFYSIGEEWERLSETFKWLW-CQCLG 971

QY 97 AGRGRISCTIANRCHGGQSYKIGDTRWRPHETGGYMLECVCLNGKGEW 146
DB 972 FGSGHFRCDSSKWCNDGNNYKIGEKWDQGE-NGQMIDCTCLNGKGEF 1020

RESULT 7
A49173

fibronectin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A49173; A31371; C60597
R;Khandjian, E.W.; Salomon, C.; Leonard, N.; Tremblay, S.; Turler, H.
Exp. Cell Res. 202, 464-470, 1992
A;Title: Fibronectin gene expression in proliferating, quiescent, and SV40-infected mouse fibroblasts
A;Reference number: A49173; MUID:93011702; PMID:1327855
A;Accession: A49173
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-103 <KHA>
A;Cross-references: UNIPROT:P11276; GB:S45680; NID:G256715; PIDN:AAB23491.1; PID:G256716
A;Note: sequential source: kidney cells
A;Note: sequence extracted from NCBI backbone (NCBIN:115080, NCBP:115081)
R;Blatti, S.P.; Foster, D.N.; Ranganathan, G.; Moses, H.L.; Getz, M.J.
Proc. Natl. Acad. Sci. U.S.A. 85, 1119-1123, 1988
A;Title: Induction of fibronectin gene transcription and mRNA is a primary response to growth factors
A;Reference number: A31371; MUID:88124987; PMID:3124113
A;Accession: A31371
A;Molecule type: mRNA
A;Residues: 1-65,'N',67-103 <BLA>
A;Cross-references: GB:M18194; GB:J03646; NID:g193329; PIDN:AAA37636.1; PID:g387159
R;Ryseck, R.P.; Macdonald-Bravo, H.; Zerial, M.; Bravo, R.
Exp. Cell Res. 180, 537-545, 1989
A;Title: Coordinate induction of fibronectin, fibronectin receptor, tropomyosin, and actin during neurite outgrowth
A;Reference number: A60597; MUID:89121031; PMID:2521606
A;Accession: C60597
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-103 <RXS>
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeats
C;Keywords: alternative splicing; duplication; extracellular matrix; glycoprotein; heterodimer
F;11-46/Domain: fibronectin type I repeat homology <1F12>
F;11-37,35-46/Disulfide bonds: #status predicted
F;84/Disulfide bonds: interchain (to 88) #status predicted
F;88/Disulfide bonds: interchain (to 84) #status predicted

Query Match 10.9%; Score 164; DB 2; Length 103;
Best Local Similarity 47.4%; Pred.No. 3.9e-06;
Matches 27; Conservative 12; Mismatches 16; Indels 2; Gaps 1;

Qy 21 CYDNGKHVQIQNWERTYLVGNLVCYCGRSGFNCS--KPAAETCFDKYTGYNTY 75
||||| : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 11 CYDDGKTYHVGEOQKEYLGAICSTCFGQGRGRCDCRRPGAAEPSPDGTGHTY 67
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 8
A56136
jagged protein precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 21-Jul-2003
C;Accession: A56136
R;Lindsell, C.E.; Shawber, C.J.; Boulter, J.; Weinmaster, G.
Cell 80, 909-917, 1995
A;Title: Jagged: a mammalian ligand that activates Notch1.
A;Reference number: A56136; MUID:95211842; PMID:7697721
A;Accession: A56136
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1220 <LIN>
A;Cross-references: GB:L38483
F;379-410/Domain: EGF homology <EGF1>
F;492-523/Domain: EGF homology <EGF>
F;634-665/Domain: EGF homology <EGF2>

```

Query Match      9.2%; Score 138; DB 2; Length 1220;
Best Local Similarity 23.9%; Pred. No. 0.0042;
Matches 70; Conservative 26; Mismatches 95; Indels 102; Gaps 19;

Qy 21 CYDNGKHYY--INQWERTY-----LGNVLVCTCYGSGRGNCE 57
Db 645 CIDGVNSYKICSDGWSGAHCENNIDCSNPCHYGGTCRDVLNDFYCDCKNGWKGTCH 704

```

Qy	58	SKPEA--EETCFDKYGTNTVRVGGTVERPKDSMIWDCTCAGRGRISCTIA-----N	108
Dd	705	SRDSQCDEATC--NNGGTCYDEVDTFK-----CMCPGSGWEG-TTCNIARNSSCLPN	752
Qy	109	RCEGGOSYKIGDTRRRPHEPTGGVMLECYCLGNKGGEWTCKPTAEKCFDHA--AGTSYV	165
Dd	753	PCHNGGTCVVNGDSF-----TCVCKEGWGP-ICTQNTNDCSHPYCNSGTC-V	799
Qy	166	VGETWEKPYO----GMMMVDC-----TCLGEGSG-RITCT---SRNRC	200
Dd	800	DGNW---YKCECAPGAPGPCRNINECOSSPCAFCATCUDELINGVQCICPPGHSGAKC	856
Qy	201	NDOOTRTSYRGDT--WSKKDNRGNNLLQCICTGNR-----GEWKGERH	242
Dd	857	HEVSGRSCTIIMGRVILDGAKWDDCNTCOCL--NGRVACSKVCGWPRLHL	906

RESULT 9
 T42215
 zonadhesin - mouse

N;Alternate names: sperm-specific membrane protein
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T42215
 R;Gao, Z.; Garbers, D.L.
 J. Biol. Chem. 273, 3415-3421, 1998

A;Title: Species diversity in the structure of zonadhesin, a sperm-specific memm
 A;Reference number: Z22080; MUID:98123114; PMID:9452463
 A;Accession: T42215
 A;Status: preliminary; translated from GE/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-5376 <NA>
 A;Cross-references: UNIPROT:O88799; EMBL:U97068; NID:g3327420; PID:g3327421; PI
 C;Genetics:
 A;Gene: Zan
 A;Map position: 5
 C;Function:
 A;Description: functions in multiple cell adhesion processes
 A;Note: found exclusively on the apical region of the sperm head
 C;Keywords: cell adhesion

Query Match	9.0%; Score 135; DB 2; Length 5376;
Best Local Similarity	22.6%; Pred. No. 0.028;
Matches	68; Conservative 34; Mismatches 105; Indels 94; Gaps 17;
Qy	14 VSOSKPGCYD-NGKHVQIQOWERTYLGVLVCTCYGGS---RFNCESKPEABETCFDK 69
Db	4557 VPSECCCKDAHGVLLPESKITW---VSRGCTYKNTCKGGTVQCHDFSPCT---GSRCLDN 4610
Qy	70 YTGNTYRVGDTYBRPKDSMIWDC--TCIGAGR-----GRISCTIANRCHEG----- 113
Db	4611 NEGNSNCVTYALKCPAHSLYTNCLPCLPSCSDPEGLCGGTSPEVPSTCKEGCICQSGYV 4670
Qy	114 -----GOSYKIGDTWRPHETGGYMLECVLGNKGGEWTK-----PI 151
Db	4671 LHKNNKMLRIHCDCKDPQGSLLIKGTQW-----ISSGGSKICTCKG---GPFQCSYKCP 4723
Qy	152 AEKCFDHAAGTSYVVGSETWPKPYGVMWVDC--TCLG-----EGS----- 189
Db	4724 GTQCEESDSSNCVSTMKCPANS-LYTHCLPTCLPSCSNPDGRCHGTSHKAPSTCREG 4782
Qy	190 -----GRITCTSNNRCDODTRTS-YRIGDTWSKKDNRGNLLQICITGNRGGEWKC 239
Db	4783 CVCQPGYLLNKDTCVHKNOGCKDIRGNIIPAGNTWISSDCTQS---CACTD---GVIOC 4836
Qy	240 E 240
Db	4837 Q 4837

RESULT 10
A56175

adhesive plaque protein Mgf2 precursor - Mediterranean mussel
C;Species: Mytilus galloprovincialis (Mediterranean mussel)
C;Date: 27-Apr-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56175
J;Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.
J. Biol. Chem. 270, 6698-6701, 1995
A;Title: Mussel adhesive plaque protein gene is a novel member of epidermal growth factor family
A;Reference number: A56175; MUID:95204464; PMID:7896812
A;Accession: A56175
A;Molecule type: mRNA
A;Residues: 1-473 <INO>
A;Cross-references: UNIPROT:Q25464; GB:D43794; NID:G602767; PIDN:BAA07852.1; PID:d100843
C;Keywords: duplication
F;1-17/Domain: signal sequence #status predicted <SIG>
F;387-419/Domain: EGF homology <EGF1>
F;429-460/Domain: EGF homology <EGF>
F;23,36,43,56,75,382,424,455,468,473/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #

Query Match 8.6%; Score 128.5; DB 2; Length 473;
Best Local Similarity 25.1%; Pred. No. 0.0096;
Matches 60; Conservative 21; Mismatches 85; Indels 73; Gaps 16;

QY 40 GNVLVCTCYGSGRGFNCEKPEAEETCFDKYTGNTYR---VGDYERPKDSMIWDCTCTIG 96
DB 64 GGSYKCFCKGGYGYGNLXK---NACKNQCKKRCVPGVTKFK-----CVCRN 110
QY 97 AGGRIS---CTIANRCHGGGYSYKIGDTWRPHETGGYMLCVCVLGNGKGB---WTC 148
DB 111 GNFRLCEKNVCS-PNPKCKNGKCSPLGKT-----GY--KCTCSGGYTGPRCEVHAC 159
QY 149 KPTAEKCFDHAAGTSYVVGTEKPYQGWMMVDTCLGEGSGRITCTSR-----NRCNDQD 204
DB 160 KP--NPCKN--KGRCPFDGKTYK-----CRCVDGSGP-TCQENACKNPNCNGG 205
QY 205 TRTSYRIGDTWSKKDNRLQICITGNGRGWKERHTSVQT-----TSSGGSPF 255
DB 206 TCSADKFGD-----YSCECRPGYFGP-ECERYVCAPNPKNGGICSDGSGGY 252

RESULT 11
A35672
crumbs protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 09-Jul-2004
C;Accession: A35672
R;Teppas, U.; Theres, C.; Knust, E.
Cell 61, 787-799, 1990
A;Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila
A;Reference number: A35672; MUID:90263104; PMID:2344615
A;Accession: A35672
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2139 <TEP>
A;Cross-references: UNIPROT:P10040; GB:M33753
A;Note: the authors translated the codon GGC for residue 1928 as Cys, and TAT for residue
C;Genetics:
A;Gene: FlyBase:Crb
A;Cross-references: FlyBase:FBgn0000368
C;Keywords: transmembrane protein
F;352-385/Domain: EGF homology <EGX1>
F;392-424/Domain: EGF homology <EGF1>
F;691-722/Domain: EGF homology <EGF>
F;767-799/Domain: EGF homology <EGF3>
F;1878-1914/Domain: EGF homology <EGX2>

Query Match 8.1%; Score 121; DB 2; Length 2139;
Best Local Similarity 23.1%; Pred. No. 0.15;
Matches 59; Conservative 22; Mismatches 88; Indels 86; Gaps 15;

QY 16 QSKPGCVGDKGHVQINQWERTYLVNVLVCTCYGSGRGFNCE-----SKP-EAETCF 67
DB 353 QTNP-CLNNGACVVIG-----SGALTCECPKYGARCEVDTDECAEQPCQNNGSCI 404

QY 68 DKYGTNTYVGDYERPKDSMIWDCTCTCAGGRISCTI-----ANRCHGGGYSYKIGD 121
DB 405 DRING-----FSCDCSGTGTGAFQCTNVNDECDKNPCLNGG-----440
QY 122 TWRRPHETGGYMLCVCVLGNGKGEWTKPI---AEKCFDHAAGTSYVVGTEKPYQGWMM 178
DB 441 --RCLHTYGYW--TCCLDGGWGEICDRPMTCTQOCFNG-----GTCLDKP-----I 484
QY 179 MVDCTCLGEGSGRITCTSRNQDQTRTSYRIGDTWSKKDNRLQICITGNGRGWK 238
DB 485 GFQCLCPPEYTGEL-CQIAPSCAQCPIDSECVGG-----KCVCKPGSSG-VN 530
QY 239 CERHTSVQTTSSGSG 253
DB 531 CQ-----TSTGDG 538

RESULT 12
A32230
tenascin precursor - chicken
N;Alternate names: cytactinin; hexabrachion
N;Contains: tenascin 190K; tenascin 200K
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A32230; B32230; A33379; B33379; C33379; S01292; A30903
R;Jones, P.S.; Hoffman, S.; Cunningham, B.A.; Edelman, G.M.
Proc. Natl. Acad. Sci. U.S.A. 86, 1905-1909, 1989
A;Title: A detailed structural model of cytactinin: protein homologues, alternative RNA
A;Reference number: A32230; MUID:89184536; PMID:2467292
A;Accession: A32230
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1810 <JON>
A;Cross-references: UNIPROT:Q90824; GB:J04519; NID:G211717; PIDN:AAA48745.1; PID:G21171
A;Accession: B32230
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1044,1318-1810 <J02>
A;Cross-references: GB:J04519
R;Spring, J.; Beck, K.; Chiquet-Ehrismann, R.
Cell 59, 325-334, 1989
A;Title: Two contrary functions of tenascin: dissection of the active sites by recombin
A;Reference number: A33379; MUID:90030407; PMID:2478295
A;Accession: A33379
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-204,'G',206-221,'A',223-380,'D',382-386,'H',388-444,'HN',447-450,'V',452-
A;Cross-references: GB:M23121; NID:G212746; PIDN:AAA49086.1; PID:G212749
A;Accession: B33379
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr
A;Molecule type: mRNA
A;Residues: 1-204,'G',206-221,'A',223-380,'D',382-386,'H',388-444,'HN',447-450,'V',452-
A;Accession: C33379
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr
A;Molecule type: mRNA
A;Residues: 1-204,'G',206-221,'A',223-380,'D',382-386,'H',388-444,'HN',447-450,'V',452-
A;Cross-references: GB:M23121
R;Pearson, C.A.; Pearson, D.; Shibahara, S.; Hofsteenge, J.; Chiquet-Ehrismann, R.
EMBO J. 7, 2977-2982, 1988
A;Title: Tenascin: cDNA cloning and induction by TGF-beta.
A;Reference number: S01292; MUID:89030589; PMID:2460335
A;Accession: S01292
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 27-181,'R',183-204,'G',206-221,'A',223-380,'D',382-386,'H',388-444,'HN',447
A;Cross-references: EMBL:X08030
A;Note: part of this sequence was confirmed by protein sequencing
C;Superfamily: tenascin; EGF homology; fibronogen beta/gamma homology; fibronectin type
C;Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracel
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-33/Domain: propeptide #status predicted <PRO>
F;34-1810/Product: tenascin 230K #status predicted <MAT>
F;223-249/Domain: EGF homology <EGF1>

F:316-342/Domain: EGF homology <EGF>
F:592-673/Domain: fibronectin type III repeat homology <FN3A>
F:681-765/Domain: fibronectin type III repeat homology <FN3B>
F:773-857/Domain: fibronectin type III repeat homology <FN3C>
F:865-949/Domain: fibronectin type III repeat homology <FN3D>
F:957-1037/Domain: fibronectin type III repeat homology <FN3E>
F:1046-1128/Domain: fibronectin type III repeat homology <FN3F>
F:1137-1219/Domain: fibronectin type III repeat homology <FN3G>
F:1228-1310/Domain: fibronectin type III repeat homology <FN3H>
F:1318-1399/Domain: fibronectin type III repeat homology <FN3I>
F:1407-1487/Domain: fibronectin type III repeat homology <FN3J>
F:1495-1575/Domain: fibronectin type III repeat homology <FN3K>
F:1590-1798/Domain: fibronectin type III repeat homology <FN3L>
F:1734-1747/Domain: calcium binding #status predicted <CAB>

Query Match 7.9%; Score 118.5; DB 1; Length 1810;
Best Local Similarity 23.1%; Pred. No. 0.2;
Matches 74; Conservative 24; Mismatches 107; Indels 115; Gaps 22;
QY 8 POSPVAVSQ--SKPGCYDNGKHVQINQWERTYLGVLVCTCYGSGRGFNCSKPEAET 65
Db PMSQTAAGRLDTAPYCSHG-----NYSTEICGCVCEPGWKGPNC--SEPACPRN 195
QY 66 CFDK-----YTGNTYR-----VGDYTERPK--DSMIWDCTCIGAGRG----- 100
Db CLNRGLCVRAKICICEBGTGDCSQARCPDNDQCKVDGV---CVCFEGYTGPDGCEE 252
QY 101 --RISCTIANR-----CHEGQSYKIGDTWRP-----HETGGYM-LRCVC-----LG 140
Db LCPHGGGTHGRCVGRVCVCHG-----FTGEDCNEPLCPNNCHNRGRCVDNECVCEGYTG 308
QY 141 NGKGWETCKPIAEKCFDHAA---GTSYVVGTEWPKYQGMVMDCTCLG--EGSGRI--- 192
Db EDCGELIC---PNDCFDRGRGCTGFC-----EGYTGDCGELTCPNNGNGRCENG 360
QY 193 -----TCTSRNRCD-----QDTRTSYRIGDTWSKD--NRGNLL 225
Db LCVCHGFGVDDCSQKRCPKTCNNRGRCVDCRGVCHEGYLGDCGELRCPNDCHNRGCI 420
QY 226 --QCIC-----TNGRGWK 239
Db NGQCVCEGFIGEDCGELRC 440

RESULT 13
T22674
hypothetical protein F54F3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22674
R:Percy, C.; Lloyd, C.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19598
A:Accession: T22674
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1584 <WIL>
A:Cross-references: UNIPROT:Q33791; EMBL:Z79696; PIDN:CAB01972.1; GSPDB:GN00023; CESP:F54F3
A:Experimental source: clone F54F3
C:Genetics:
A:Gene: CESP:F54F3.1
A:Map position: 5
A:Introns: 35/3; 157/3; 213/2; 257/3; 357/1; 440/3; 545/1; 854/3; 896/1; 944/1; 1001/1;

Query Match 7.8%; Score 116.5; DB 2; Length 1584;
Best Local Similarity 21.8%; Pred. No. 0.26;
Matches 61; Conservative 21; Mismatches 123; Indels 75; Gaps 17;
QY 5 MVQPSPVAVSQSKPCYDNGKHVQINQWERTYLGVLVCTCYGSGRG---FNC-----E 57
Db 1058 MVPAQPKTCVSS--DCHING-HCVINEHGAGEY-----ICQCLPGFSGDGFNCRGADQ 1109
QY 58 SKPEAETCFDKYTGNTYRIGDTWRPDKSMIWDCTCIGAGRG-----RISCTTIA 107

Db 1110 CNPNSPSACYQ-----NAHCVVYDAILNAH-----ACKCVDFGKGDGTSVCVPYAPATNCNLE 1160
QY 108 NR-CHEGQSYKIGDTWRPHEGMYLCEVCLGNKGKEWTCK-----PIAEKCFDHAA 160
Db PRICHANAQCV-----MHEDTNAY--ECICKPGSGDGYTKCDVIEPTRCTNSIHA- 1210
QY 161 GTSYVVGTEWPKYQGMVMDCTCIGESGRITCTSRNRCDNDQDTRTSYRIGDTWSKKN 220
Db 1211 ---YCA-----QNPTSGAYQCKNAGYNGNGHL-CVSMSSCLDDR-----SLCDE 1251
QY 221 RGNLLQICICTNGRGWKRCERTHSYQTTSSSGSPFTDVRA 260
Db 1252 NAD-----CVPGEAGHYVCNCHYGVHGDGRSCSPESSTRS 1286

RESULT 14
T15099
hypothetical protein W03F8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15099
R:Johnson, D.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid W03F8.
A:Reference number: Z18293
A:Accession: T15099
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1808 <JOH>
A:Cross-references: UNIPROT:O44565; EMBL:AF039041; NID:g2736380; PID:g2736388; PIDN:AAAB9
A:Experimental source: strain Bristol N2; clone W03F8
C:Genetics:
A:Gene: CESP:W03F8.5
A:Map position: 4
A:Introns: 99/2; 255/2; 365/2; 435/3; 565/2; 975/2; 1055/2; 1602/3; 1719/3; 1766/1
C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
F:797-842/Domain: laminin-type EGF-like homology <LEG>

Query Match 7.7%; Score 115; DB 2; Length 1808;
Best Local Similarity 20.9%; Pred. No. 0.38;
Matches 62; Conservative 26; Mismatches 107; Indels 102; Gaps 14;
QY 8 POSPVAVSQKPGCY--DNGKHVQINQWERTYLGVLVCTCYGSGRGFNCSKPEAET 65
Db PGGTSGYQHADTCYLRSNNQ-----DIVCNCKSYQGERGCE----- 985
QY 66 CFDKYTGNTYRIGDTWRPDKSMIWDCTCIGAGRGRICTTIANRCHGEGQSYKIGDTWR 125
Db CAQNHGSPREVGGTCER-----CDC---NGNIDWAM-----EGSCDAATGCECLKC 1028
QY 126 PHETGGYMLE-----CVC--LNGKGWETCKPIAEKCFDHAAAGTSYV 166
Db 1029 LHHTEGAQCEHCVDGYGDAKLKTCQRCVCELNELGNTSTKGACDRVSGQCPCF-----DNII 1084
QY 167 GETWEKPYQGMVMDCTC-----LGEESGRITC-----TSRNRCDQDTRT 207
Db 1085 G-----MQCD-QCAENHFNLASGACBACCDSNGVVLNHEGVPHLQCNIFDGQC 1133
QY 208 SYRIGDTWSKKNRGNLL-----QCICTNGRGWKRCERTHSYQTTSSSG 253
Db 1134 QCKPGRGKCKQCDLWGDPTTDPDCHRCBCNPTGSKSLQCHRNNGTCECQAGSG 1190

RESULT 15
MMFFBI
laminin beta-1 chain precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: laminin chain B1
C:Species: Drosophila melanogaster
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C:Accession: A28783; S14462; B28783
R:Montell, D.J.; Goodman, C.S.
Cell 53, 463-473, 1988


```
Db      1013 ACLKLYQT 1022

Search completed: September 30, 2005, 13:31:49
Job time : 42.4373 secs
```

Query Match	7.6%;	Score 114;	DB 1;	Length 1790;
Best Local Similarity	24.4%;	Pred. No. 0.45;		
Matches	61;	Conservative 18;	Mismatches 97;	Indels 74; Gaps 15;
Qy	44	VCTCYGSGRGFNCESKPEAEETCFDKYGTNTYRVGDTYERPXSMTWDCICIGAG-----	98	
Db	802	VCESNGG----YCOCKPNVVRQCDQCAPGTGYFG-----PEGKACDCNCSIGSKDYKD	852	
Qy	99	--RGRISC---TIANRCHEGSQSYKIGDTWRRPHETGGYMLEC-VCLGNKGGEWTKCKPIA	152	
Db	853	LITGQCQCVENTYRECNCQCPGY-----WNFP-----ECFVCCQGHAA-TCDDPIQ	898	
Qy	153	EKCPDHAAGTSYVVGETWTEPKYQGMVM-----DCTC-----LGECSGRITCTSRN	198	
Db	899	GTICIDCQDSTTGYSCSLGYYGNPLFGSEIGCRPCRPETVASGHLAHADGCSLDTRNN	958	
Qy	199	R--CNDOD---TFTSYRIGDTWSKKDNRNLLQIC-----TCNGRGEWKCEBHTS	244	
Db	959	NMLCHCEGYSGRKCEICADNFPNDNGGTCSKCECSNNVDLYDTGN-----CDKQTG	1012	
Qy	245	V-----QTT	248	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 08:57:15 : Search time 33.6978 Seconds
(without alignments)
580.396 Million cell updates/sec

Title: US-09-492-971b-15_COPY_4_265
Perfect score: 1499
Sequence: 1 QAQMVPQSPVAVSQSKPG.....TSVQTSSGSGPTDYRAAV 262

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499	100.0	2231	1	Sequence 16, Appl
2	1499	100.0	2324	1	Sequence 1, Appli
3	1499	100.0	2324	5	Sequence 1, Appli
4	1499	100.0	2386	2	Sequence 12, Appl
5	1499	100.0	2446	2	Sequence 2, Appli
6	1499	100.0	2446	5	Sequence 2, Appli
7	1495	99.7	2386	4	Sequence 1, Appli
8	1494	99.7	2327	6	Patent No. 5455158
9	1494	99.7	2327	6	Patent No. 5455158
10	295.5	19.7	474	2	Sequence 9, Appli
11	295	19.7	422	2	Sequence 12, Appl
12	294.5	19.6	188	1	Sequence 14, Appl
13	159	10.6	1193	2	Sequence 10, Appl
14	159	10.6	1193	3	Sequence 10, Appl
15	159	10.6	1193	4	Sequence 10, Appl
16	148	9.9	1010	3	Sequence 7, Appli
17	148	9.9	1010	4	Sequence 7, Appli
18	148	9.9	1036	3	Sequence 6, Appli
19	148	9.9	1067	4	Sequence 18, Appl
20	148	9.9	1187	3	Sequence 7, Appli
21	148	9.9	1208	4	Sequence 1, Appli
22	148	9.9	1208	4	Sequence 1, Appli
23	148	9.9	1218	2	Sequence 6, Appli
24	148	9.9	1218	3	Sequence 6, Appli
25	148	9.9	1218	3	Sequence 2, Appli
26	148	9.9	1218	3	Sequence 7, Appli
27	148	9.9	1218	3	Sequence 11, Appl

Sequence 7, Appli
Sequence 2, Appli
Sequence 85, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 5902, Ap
Sequence 10297, A
Sequence 5, Appli
Sequence 5, Appli
Sequence 20, Appl
Sequence 17, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 6, Appli
Sequence 4, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US 07/847975
; APPLICATION NUMBER: 06-MAR-1992
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..2231
; OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match 100.0%; Score 1499; DB 1; Length 2231;
Best Local Similarity 100.0%; Pred. No. 1.9e-136;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 60
DB 1 QAQMVPQSPVAVSQSKPCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 60
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTIANRCHGGQSYKIG 120
DB 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTIANRCHGGQSYKIG 120
QY 121 DTWRRPHETGGYMLCVCCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMVMV 180
DB 121 DTWRRPHETGGYMLCVCCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMVMV 180
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNGRGWKCE 240
DB 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNGRGWKCE 240
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 241 RHTSVQTTSSGSGPFTDVRAAV 262

RESULT 2
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,857
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..2231
; OTHER INFORMATION: /note= "Human fibronectin"
US-08-283-857-1

Query Match 100.0%; Score 1499; DB 1; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2e-136;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 60
DB 1 QAQMVPQSPVAVSQSKPCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCEKSP 60
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTIANRCHGGQSYKIG 120
DB 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTIANRCHGGQSYKIG 120
QY 121 DTWRRPHETGGYMLCVCCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMVMV 180
DB 121 DTWRRPHETGGYMLCVCCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMVMV 180
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNGRGWKCE 240
DB 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNGRGWKCE 240
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 241 RHTSVQTTSSGSGPFTDVRAAV 262

RESULT 3
PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match      100.0%; Score 1499; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2e-136;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQVQPQSPVAVSQKPGCYDNGKHQYINQOQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 1 QAOQVQPQSPVAVSQKPGCYDNGKHQYINQOQWERTYLGVLVCTCYGSGRGFNCESKP 60
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGQSYKIG 120
Db 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGQSYKIG 120
Qy 121 DTWRPHETGGYMLECVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEKPYQGMMV 180
Db 121 DTWRPHETGGYMLECVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEKPYQGMMV 180
Qy 181 DCTCLGSGSGRICTSRNRNCNDQDTRTSYRIGDTWSKKDNKRNLLQICITGNGRGWKCE 240
Db 181 DCTCLGSGSGRICTSRNRNCNDQDTRTSYRIGDTWSKKDNKRNLLQICITGNGRGWKCE 240
Qy 241 RHTSVQTTSSGSGPFTDVRVAV 262
Db 241 RHTSVQTTSSGSGPFTDVRVAV 262

RESULT 4
US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-09-016-366A-12
Query Match      100.0%; Score 1499; DB 2; Length 2386;
Best Local Similarity 100.0%; Pred. No. 2e-136;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQVQPQSPVAVSQKPGCYDNGKHQYINQOQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 32 QAOQVQPQSPVAVSQKPGCYDNGKHQYINQOQWERTYLGVLVCTCYGSGRGFNCESKP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGQSYKIG 120
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGQSYKIG 151
Qy 121 DTWRPHETGGYMLECVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEKPYQGMMV 180
Db 152 DTWRPHETGGYMLECVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEKPYQGMMV 211
Qy 181 DCTCLGSGSGRICTSRNRNCNDQDTRTSYRIGDTWSKKDNKRNLLQICITGNGRGWKCE 240
Db 212 DCTCLGSGSGRICTSRNRNCNDQDTRTSYRIGDTWSKKDNKRNLLQICITGNGRGWKCE 271
Qy 241 RHTSVQTTSSGSGPFTDVRVAV 262
Db 272 RHTSVQTTSSGSGPFTDVRVAV 293

RESULT 5
US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246
; REFERENCE/DOCKET NUMBER: 92-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-356-2
Query Match      100.0%; Score 1499; DB 2; Length 2446;
```

Best Local Similarity 100.0%; Pred. No. 2.1e-136;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
DB 32 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGNFNCESKP 91

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151

QY 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
DB 152 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 211

QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRNLLQICITGNRGRGWKCE 240
DB 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRNLLQICITGNRGRGWKCE 271

QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 6
PCT-US93-12687-2
; Sequence 2, Application PC/TUS9312687
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher H.
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12687
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/998,271
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 92-26PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-12687-2

Query Match 100.0%; Score 1499; DB 5; Length 2446;
Best Local Similarity 100.0%; Pred. No. 2.1e-136;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
DB 32 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGNFNCESKP 91

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151

QY 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
DB 152 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 211

QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRNLLQICITGNRGRGWKCE 240
DB 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRNLLQICITGNRGRGWKCE 271

QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 7
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match 99.7%; Score 1495; DB 4; Length 2386;
Best Local Similarity 99.6%; Pred. No. 5e-136;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
DB 32 QAOQWVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGVLVCTCYGSGRGNFNCESKP 91

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151

QY 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 180
DB 152 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMWV 211

QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRNLLQICITGNRGRGWKCE 240
DB 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKDNRNLLQICITGNRGRGWKCE 271

QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 8
5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA, LEVANON, AVIGDOR, WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME

RESULT 10
 US-08-836-854-9
 ; Sequence 9, Application US/08836854
 ; Patent No. 5824547
 ; GENERAL INFORMATION:
 ; APPLICANT: HASHINO, Kimikazu
 ; APPLICANT: MATSUSHITA, Hideyuki
 ; APPLICANT: KATO, Ikumoshin
 ; TITLE OF INVENTION: METHOD OF PRODUCTION OF TRANSFECTED CELLS
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street N.W. Ste. 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/836.854

PRIOR APPLICATION DATA: PCT/JP95/02425
 APPLICATION NUMBER: PCT/JP95/02425
 FILING DATE: 29-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 317721/1994
 FILING DATE: 29-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25,618
 REFERENCE/DOCKET NUMBER: HASHINO=1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 474 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 PS-08-836-854-9

	Query Match	19.7%;	Score 295.5;	DB 2;	Length 474;
	Best Local Similarity	33.8%;	Pred. No. 2.1e-20;		
	Matches	72;	Conservative	22;	Mismatches 70; Indels 49; Gaps 9
Qy	3	QQWVQP--QSPVAVSQSKPGCYDNGKHVQIQINQOWERTYLGNVLVCTCVGSGRGFNCSKP	60		
Db	227	QEFTVPGSKSTATISGLKPGV-----DYTTIVYAVTGRGDS---PASSKP	268		
Qy	61	-----EABETCFDKYTGTNTYRVGDTYVERPKDS---MIWDCTCIGA	97		
Db	269	ISINYRTEIDKPSONEGINOPTDDSCFDPTTVSHYAVGWDERMSEGSFKLL--COCLGF	326		

98	QY	GGGRISCTIANRCHGGGGSYKIGDTWRPHETGTGMYLCEVCLNGNGKGEWTKCIAEKQFD	157
327	Db	GSGHFRCDSSRWCHDNGVNYKIGEKWDQGB-NQOMMSCTCLNGNGKGFKCDPHEATCYD	385
158	QY	HAAGTSYVVGETWEPYQGMWMDCTCLGEGSG	190
386	Db	D--GKTYHVGEOWKEYLK-AICSTCTFGGORG	415

RESULT 11
US-08-836-854-12
; Sequence 12, Application US/08936854
; Patent No. 5824547
; GENERAL INFORMATION:
; APPLICANT: HASHINO, Kimikazu
; APPLICANT: MATSUSHITA, Hideyuki
; APPLICANT: KATO, Ikuneshin
; TITLE OF INVENTION: METHOD OF PRODUCTION OF TRANSFECTED CELLS

Query Match	19.7%	Score 295;	DB 2;	Length 422;
Best Local Similarity	33.6%	Pred. No. 2e-20;		
Matches	72;	Conservative 22;	Mismatches 70;	Indels 50; Gaps 9;

Qy	3	QOMVQP--QSPVAVSQSPGCGYDNGKHVQIQOWERTYLGNVLVCTCYGSGRGFNCSEKSP	60
	:	: : : : :	
Db	223	QEFTVPGSKSPATISGLKPGV-----DYITIVYAVTGRGDS---PASSKP	264
	:	: : : : :	
Qy	61	-----EAEETCPDKYTGNTYRGVDYERPKDS---MWDCTCIG	96
	:	: : : : :	
Db	265	ISINRYTEIDKPSMANEGLNQPTDDSCDFPTVSHYAVGDSEWRNMSGFKLL--CQCIG	322
	:	: : : : :	
Qy	97	AGGRISCTIANRCHEGSQSYKIGDTWRRRPHETGYLMCEVCLGNGKEWTCCKPIAEKCF	156
	:	: : : : :	
Db	123	FGSGHERCDSRWCHDGVNVIKIGKWDQCE--NGOMMSTCTCINGKGFCKDCDHEATCY	381
	:	: : : : :	

```

QY      157 DHAACTSYVVGTEWXPYOGHMMVVDCTCLGEGSG 190
      |   |   |   |   |   |   |   |   |   |   |
      |   |   |   |   |   |   |   |   |   |   |
Db      382 DD--GKTYHVGQWQKEYLG-AICSTCFGGQRG 412
      |   |   |   |   |   |   |   |   |   |   |
      |   |   |   |   |   |   |   |   |   |   |

RESULT 12
; Sequence 14, Application US/08142449B
; Patent No. 5668104
; GENERAL INFORMATION:
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kawano, Genji
; APPLICANT: Sudo, Tetsuo
; APPLICANT: Kojima, Katsuaki
; TITLE OF INVENTION: Physiologically Active Protein and
; TITLE OF INVENTION: Hematopoietic Stem Cell Growth Agent
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nels T. Lippert, White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,449B
; FILING DATE: 24-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lippert, Nels T.
; REGISTRATION NUMBER: 25,888
; REFERENCE/DOCKET NUMBER: 1145358-304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)-819-8582
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-142-449B-14

```

```

Query Match          19.6%; Score 294.5; DB 1; Length 188;
Best Local Similarity 43.5%; Pred. No. 8e-21;
Matches 57; Conservative 19; Mismatches 46; Indels 9; Gaps 5;

QY      63  EETCFDYKNTYRVGDTYBRPKDS---MTWDDTCICAGRGRICTIANRCHGGQSYKI 119
      :::||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||
Db       5  DDCSFDYTYVSHYAVGVDEWERMESGKLLI--COCFLFGSGHFRCDSSRWCHDGNVYKI 62
      :::||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||

QY      120 GDTWRRRHETGGYMLECVCLGNLNGKEWTCKPIAEKCFDHAAGTSYVVVGETWEKPYQGMM 179
      :::||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||
Db       63  GEKWDRQGE-NGQWMSCTCLGNLNGKEPKCDPHEATCYDD--GKTYHVGQWQKEYLGI-AI 118
      :::||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||

QY      180 VDCTCLGEGSG 190
      |||||
Db      119  CSCTCFGGQRG 129
      |||||

RESULT 13
US-08-400-159-10
; Sequence 10, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.

```


APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,159
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1193 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-400-159-10

Query Match 10.6%; Score 159; DB 2; Length 1193;
Best Local Similarity 26.4%; Pred. No. 1.2e-06;
Matches 73; Conservative 22; Mismatches 88; Indels 94; Gaps 18;

Qy 21 CYDNGKHQY--INQWERTY-----LGNVLVCTCYGSGRGFNC 57
Db 618 CIDGVNSYKICSDGEGTYCETINDCSKNPCHNGGTCDLVNDFCECKNGWKGTCH 677

Qy 58 SKPEA--EETCFDKYTGNTYRVGDYTPKDSMIWDCITCIGAGRGISCTIA-----N 108
Db 678 SRDSQCDEATC--NNGGTCYDEGDTFK-----CMC-PAGWEGATCNIARNSSCLPN 725

Qy 109 RCHEGGSYKIGDTWRPHETGGYMLBCVCLGNGKEWTCKPIAEKCFDHA---AGTSYV 165
Db 726 PCHNGGTCVVGDSF-----TCVCKEGWGP-TCTQNTNDCSPHPCTNSGTC-V 772

Qy 166 VGETWEKPYQ-----GMMVMDC-----TCLGEGSG-RITC---TSNRNC 200
Db 773 DGDNW---YRCEAPGAPGPDCHININECQSSPCAFGATCVDINGYRCICPPGRSGPGC 829

Qy 201 NQDQTR---TSYRIGDTWSKKDNRGNLLQICITGNR 234
Db 830 QEVTRGTCFTSIRVMPDGAKWDDDCNTCQCL---NGK 863

RESULT 14
US-08-611-729A-10
; Sequence 10, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.

APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,729A
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1193 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-611-729A-10

Query Match 10.6%; Score 159; DB 3; Length 1193;
Best Local Similarity 26.4%; Pred. No. 1.2e-06;
Matches 73; Conservative 22; Mismatches 88; Indels 94; Gaps 18;

Qy 21 CYDNGKHQY--INQWERTY-----LGNVLVCTCYGSGRGFNC 57
Db 618 CIDGVNSYKICSDGEGTYCETINDCSKNPCHNGGTCDLVNDFCECKNGWKGTCH 677

Qy 58 SKPEA--EETCFDKYTGNTYRVGDYTPKDSMIWDCITCIGAGRGISCTIA-----N 108
Db 678 SRDSQCDEATC--NNGGTCYDEGDTFK-----CMC-PAGWEGATCNIARNSSCLPN 725

Qy 109 RCHEGGSYKIGDTWRPHETGGYMLBCVCLGNGKEWTCKPIAEKCFDHA---AGTSYV 165
Db 726 PCHNGGTCVVGDSF-----TCVCKEGWGP-TCTQNTNDCSPHPCTNSGTC-V 772

Qy 166 VGETWEKPYQ-----GMMVMDC-----TCLGEGSG-RITC---TSNRNC 200
Db 773 DGDNW---YRCEAPGAPGPDCHININECQSSPCAFGATCVDINGYRCICPPGRSGPGC 829

Qy 201 NQDQTR---TSYRIGDTWSKKDNRGNLLQICITGNR 234
Db 830 QEVTRGTCFTSIRVMPDGAKWDDDCNTCQCL---NGK 863

RESULT 15
US-09-195-524-10
; Sequence 10, Application US/09195524
; Patent No. 6703489
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.

;; APPLICANT: Lewis, Julian H.
;; APPLICANT: Myat, Anna M.
;; APPLICANT: Fleming, Robert J.
;; APPLICANT: Artavanis-Tsakonas, Spyridon
;; APPLICANT: Mann, Robert S.
;; APPLICANT: Gray, Grace E.
;; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
;; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/195,524
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/611,729
;; FILING DATE: 06-MAR-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Misrock, S. Leslie
;; REGISTRATION NUMBER: 18,872
;; REFERENCE/DOCKET NUMBER: 7326-037
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-9741/8864
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1193 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; US-09-195-524-10

Query Match 10.6%; Score 159; DB 4; Length 1193;
Best Local Similarity 26.4%; Pred. No. 1.2e-06;
Matches 73; Conservative 22; Mismatches 88; Indels 94; Gaps 18;
Qy 21 CYDNGKHQ--INQWERTY-----LGNVLVCTCYGSGRGFNC 57
Db 618 CIDGVNSYKICSDGWEGTYCETNINDCKNPCHNGGTCRDLVNDFFCECKNGWKXKTH 677
Qy 58 SKPEA--EETCFDKYTGNTYRVDYRPRKDSMIWDCTCIGAGRISCTIA-----N 108
Db 678 SRDSQCDEATC--NNGGTCYDEGDTFK-----CMC-PAGWEGATCNIA RNSSCLPN 725
Qy 109 RCHGGGQSYKIGDTWRPHETGGVLMLECVLGNKGKGTCKPIAEKCFDHA---AGTSYV 165
Db 726 PCHNGGTCVVSQGSF-----TCVCKEGWEGP-TCTQNTNDCSPHPCYNSGTC-V 772
Qy 166 VGETWEKPYQ-----GMMVDC-----TCLGSGG-RITC---TSRNR 200
Db 773 DGNW---YRCEAPGFPAGPDCRININECOSSPCAFGATCVDEINGYRCICPPRSGPGC 829
Qy 201 NDQDTR---TSYRIGDTWSKDNRGNLQICITGNR 234
Db 830 QEVTRPCFTSIRWPDGAKWDDDCNTCQCL---NGK 863

Search completed: September 30, 2005, 08:58:04
Job time : 34.6978 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2005, 08:57:20 ; Search time 197.974 Seconds
(without alignments)
549.366 Million cell updates/sec

Title: US-09-492-971B-15_COPY_4_265

Perfect score: 1499

Sequence: 1 QAOQWVQSPVAVSQSKPG.....TSVQTSSGSGPFDVRAAV 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1499	100.0	2328	14 US-10-171-311-64	Sequence 64, Appl
2	1499	100.0	2328	15 US-10-236-031B-70	Sequence 70, Appl
3	1499	100.0	2328	15 US-10-374-979-98	Sequence 98, Appl
4	1499	100.0	2328	15 US-10-182-936A-98	Sequence 98, Appl
5	1499	100.0	2328	16 US-10-477-238A-677	Sequence 677, App
6	1499	100.0	2328	16 US-10-680-287A-677	Sequence 677, App
7	1499	100.0	2328	17 US-10-477-173-677	Sequence 677, App
8	1499	99.7	463	15 US-10-144-194A-52	Sequence 52, Appl
9	1495	99.7	463	16 US-10-491-566-52	Sequence 52, Appl
10	1495	99.7	642	16 US-10-741-601-354	Sequence 354, Appl
11	1495	99.7	642	17 US-10-741-600-1066	Sequence 1066, Ap

Query Match 100.0%; Score 1499; DB 14; Length 2328;

12	1495	99.7	657	16	US-10-741-601-359	Sequence 359, App
13	1495	99.7	657	17	US-10-741-600-1072	Sequence 1072, Ap
14	1495	99.7	984	16	US-10-741-601-356	Sequence 356, App
15	1495	99.7	984	17	US-10-741-600-1069	Sequence 1069, Ap
16	1495	99.7	2220	15	US-10-236-392-4	Sequence 4, Appli
17	1495	99.7	2296	16	US-10-741-601-363	Sequence 363, App
18	1495	99.7	2296	17	US-10-741-600-1075	Sequence 1075, Ap
19	1495	99.7	2355	15	US-10-144-194A-104	Sequence 104, App
20	1495	99.7	2355	15	US-10-360-101-235	Sequence 235, App
21	1495	99.7	2355	15	US-10-447-161-3	Sequence 3, Appli
22	1495	99.7	2355	16	US-10-734-564-94	Sequence 94, Appl
23	1495	99.7	2355	16	US-10-741-601-357	Sequence 357, App
24	1495	99.7	2355	16	US-10-741-601-366	Sequence 366, App
25	1495	99.7	2355	16	US-10-491-566-104	Sequence 104, App
26	1495	99.7	2355	17	US-10-741-600-1067	Sequence 1067, App
27	1495	99.7	2355	17	US-10-741-600-1078	Sequence 1078, App
28	1495	99.7	2355	17	US-10-852-335A-147	Sequence 147, App
29	1495	99.7	2355	18	US-10-287-436A-436	Sequence 436, App
30	1495	99.7	2355	18	US-10-287-436A-1137	Sequence 1137, Ap
31	1495	99.7	2355	20	US-11-040-130-28	Sequence 28, Appl
32	1495	99.7	2386	10	US-09-961-403-1	Sequence 1, Appli
33	1495	99.7	2386	16	US-10-741-601-360	Sequence 360, App
34	1495	99.7	2386	16	US-10-788-792-206	Sequence 206, App
35	1495	99.7	2386	16	US-10-618-281-32	Sequence 32, Appl
36	1495	99.7	2386	17	US-10-741-600-1071	Sequence 1071, Ap
37	1495	99.7	2386	17	US-10-868-577A-59	Sequence 59, Appl
38	1495	99.7	2386	18	US-10-485-758-4	Sequence 4, Appli
39	1495	99.7	2386	18	US-10-485-758-9	Sequence 9, Appli
40	1487	99.2	259	10	US-09-940-235-4	Sequence 2, Appli
41	1480	98.7	2320	15	US-10-236-392-2	Sequence 8, Appli
42	1476	98.5	2320	14	US-10-279-733-8	Sequence 2, Appli
43	1457	97.2	2474	18	US-10-450-763-52638	Sequence 52638, A
44	1192	79.5	1173	18	US-10-450-763-52634	Sequence 52634, A
45	888.5	59.3	231	9	US-09-925-302-548	Sequence 548, App

ALIGNMENTS

RESULT 1

US-10-171-311-64
; Sequence 64, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MKI-035
; CURRENT FILING DATE: 2002-06-12
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-171-311-64
Query Match 100.0%; Score 1499; DB 14; Length 2328;

Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQVQPSFVAVSQSPGCGYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
Db 5 QAOQVQPSFVAVSQSPGCGYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCSKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRISCTIANRCHGGQSYKIG 120
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRISCTIANRCHGGQSYKIG 124

Qy 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMMV 180
Db 125 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMMV 184

Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGWKE 240
Db 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGWKE 244

Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 245 RHTSVQTTSSGSGPFTDVRAAV 266

RESULT 2

US-10-236-031B-70
; Sequence 70, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236.031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-70

Query Match 100.0%; Score 1499; DB 15; Length 2328;
Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQVQPSFVAVSQSPGCGYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
Db 5 QAOQVQPSFVAVSQSPGCGYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCSKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRISCTIANRCHGGQSYKIG 120
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRISCTIANRCHGGQSYKIG 124

Qy 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMMV 180
Db 125 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMMV 184

Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGWKE 240
Db 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGWKE 244

Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 245 RHTSVQTTSSGSGPFTDVRAAV 266

RESULT 3

US-10-374-979-98
; Sequence 98, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-98

Query Match 100.0%; Score 1499; DB 15; Length 2328;
Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQVQPSFVAVSQSPGCGYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
Db 5 QAOQVQPSFVAVSQSPGCGYDNGKHQYINQOWERTYLGVLVCTCYGSGRGFNCSKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRISCTIANRCHGGQSYKIG 120
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRISCTIANRCHGGQSYKIG 124

Qy 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMMV 180
Db 125 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMMV 184

Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGWKE 240
Db 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGWKE 244

Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 245 RHTSVQTTSSGSGPFTDVRAAV 266

RESULT 4

US-10-182-936A-98
; Sequence 98, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058

```
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-98

Query Match      100.0%; Score 1499; DB 15; Length 2328;
Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCESKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRGRIISCTIANRCHGGQSYKIG 124

Qy 121 DTWRPHETGGMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGWMV 180
Db 125 DTWRPHETGGMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGWMV 184

Qy 181 DCTCLGSGGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICICTGNRGWKCE 240
Db 185 DCTCLGSGGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICICTGNRGWKCE 244

Qy 241 RHTSVQTTSSGSGPFTDVA 262
Db 245 RHTSVQTTSSGSGPFTDVA 266

RESULT 5
US-10-477-238A-677
; Sequence 677, Application US/10477238A
; Publication No. US20040221326A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-238A-677

Query Match      100.0%; Score 1499; DB 16; Length 2328;
Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCESKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRGRIISCTIANRCHGGQSYKIG 124

Qy 121 DTWRPHETGGMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGWMV 180
Db 125 DTWRPHETGGMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGWMV 184

Qy 181 DCTCLGSGGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICICTGNRGWKCE 240
Db 185 DCTCLGSGGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICICTGNRGWKCE 244

Qy 241 RHTSVQTTSSGSGPFTDVA 262
Db 245 RHTSVQTTSSGSGPFTDVA 266

RESULT 6
US-10-680-287A-677
; Sequence 677, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; PRIOR FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-680-287A-677

Query Match      100.0%; Score 1499; DB 16; Length 2328;
Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSPVAVSQSPGCGYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db 5 QAOQMVQPSPVAVSQSPGCGYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCESKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCICIGAGRGRIISCTIANRCHGGQSYKIG 124

Qy 121 DTWRPHETGGMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGWMV 180
Db 125 DTWRPHETGGMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGWMV 184

Qy 181 DCTCLGSGGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICICTGNRGWKCE 240
Db 185 DCTCLGSGGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICICTGNRGWKCE 244

Qy 241 RHTSVQTTSSGSGPFTDVA 262
Db 245 RHTSVQTTSSGSGPFTDVA 266
```

```
RESULT 7
US-10-477-173-677
; Sequence 677, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-173-677

Query Match      100.0%; Score 1499; DB 17; Length 2328;
Best Local Similarity 100.0%; Pred. No. 7.9e-130;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
DB 5 QAOQWVQPSVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYGSGRGFNCSKP 64
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 124
QY 121 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMVMV 180
DB 125 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMVMV 184
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGEWKCE 240
DB 185 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGEWKCE 244
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 245 RHTSVQTTSSGSGPFTDVRAAV 266

RESULT 8
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match      99.7%; Score 1495; DB 16; Length 463;
Best Local Similarity 99.6%; Pred. No. 3e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
DB 68 QAOQWVQPSVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYGSGRGFNCSKP 127
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 128 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 187
QY 121 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMVMV 180
DB 188 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMVMV 247
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGEWKCE 240
DB 248 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGEWKCE 307
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 308 RHTSVQTTSSGSGPFTDVRAAV 329

RESULT 9
US-10-491-566-52
; Sequence 52, Application US/10491566
; Publication No. US20040249144A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/491,566
; CURRENT FILING DATE: 2004-04-05
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-491-566-52

Query Match      99.7%; Score 1495; DB 16; Length 463;
Best Local Similarity 99.6%; Pred. No. 3e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
DB 68 QAOQWVQPSVAVSQSKPGCYDNGKHQYINQOWERTYLGNVLVCTCYGSGRGFNCSKP 127
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
DB 128 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 187
QY 121 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMVMV 180
DB 188 DTWRRPHETGGYMLECVCLGNGKGWTCCKPIAEKCFDHAAGTSYVVGGETWEKPYQGMVMV 247
QY 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGEWKCE 240
DB 248 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQICITGNRGEWKCE 307
QY 241 RHTSVQTTSSGSGPFTDVRAAV 262
DB 308 RHTSVQTTSSGSGPFTDVRAAV 329

RESULT 10
US-10-741-601-354
; Sequence 354, Application US/10741601
```

```
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match      99.7%; Score 1495; DB 16; Length 642;
Best Local Similarity 99.6%; Pred. No. 4.4e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 32 QAOQVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 91
Qy 61 EAETCFDKYTGTNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
Db 92 EAETCFDKYTGTNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 180
Db 152 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 211
Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNGRGWKCE 240
Db 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNGRGWKCE 271
Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 11
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match      99.7%; Score 1495; DB 17; Length 642;
Best Local Similarity 99.6%; Pred. No. 4.4e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 32 QAOQVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 91
Qy 61 EAETCFDKYTGTNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
Db 92 EAETCFDKYTGTNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 180
```

```
|||||
Db 152 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 211
Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNGRGWKCE 240
|||||
Db 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNGRGWKCE 271
Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 12
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match      99.7%; Score 1495; DB 16; Length 657;
Best Local Similarity 99.6%; Pred. No. 4.5e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 32 QAOQVQPQSPVAVSQKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 91
Qy 61 EAETCFDKYTGTNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 120
Db 92 EAETCFDKYTGTNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTIANRCHGGQSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 180
Db 152 DTWRRPHETGGYMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGETWEKPYQGMMV 211
Qy 181 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNGRGWKCE 240
Db 212 DCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNGRGWKCE 271
Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293

RESULT 13
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRN
; ORGANISM: Homo sapiens
```

US-10-741-600-1072

```
Query Match          99.7%; Score 1495; DB 17; Length 657;
Best Local Similarity 99.6%; Pred. No. 4.5e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGWLVCTCYGSGRGFNCEKSP 60
Db 32 QAOQMVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGWLVCTCYGSGRGFNCEKSP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLGNGKGWTKCPIAEKCFDHAAGTSYVVGWETWEKPYQGMMV 180
Db 152 DTWRRPHETGGYMLECVCLGNGKGWTKCPIAEKCFDHAAGTSYVVGWETWEKPYQGMMV 211
Qy 181 DCTCLGEGSGRIITCTSRNRCNDQDTRTSYRIGDTWSKDNRGNLLQICITGNGRGWKE 240
Db 212 DCTCLGEGSGRIITCTSRNRCNDQDTRTSYRIGDTWSKDNRGNLLQICITGNGRGWKE 271
Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293
```

RESULT 14

US-10-741-601-356

```
; Sequence 356, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-356
```

```
Query Match          99.7%; Score 1495; DB 16; Length 984;
Best Local Similarity 99.6%; Pred. No. 7.1e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGWLVCTCYGSGRGFNCEKSP 60
Db 32 QAOQMVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGWLVCTCYGSGRGFNCEKSP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLGNGKGWTKCPIAEKCFDHAAGTSYVVGWETWEKPYQGMMV 180
Db 152 DTWRRPHETGGYMLECVCLGNGKGWTKCPIAEKCFDHAAGTSYVVGWETWEKPYQGMMV 211
Qy 181 DCTCLGEGSGRIITCTSRNRCNDQDTRTSYRIGDTWSKDNRGNLLQICITGNGRGWKE 240
Db 212 DCTCLGEGSGRIITCTSRNRCNDQDTRTSYRIGDTWSKDNRGNLLQICITGNGRGWKE 271
Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293
```

RESULT 15

US-10-741-600-1069

```
; Sequence 1069, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1069
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1069
```

```
Query Match          99.7%; Score 1495; DB 17; Length 984;
Best Local Similarity 99.6%; Pred. No. 7.1e-130;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGWLVCTCYGSGRGFNCEKSP 60
Db 32 QAOQMVQPSFVAVSQSKPGCYDNGKHQYINQOWERTYLGWLVCTCYGSGRGFNCEKSP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 120
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIG 151
Qy 121 DTWRRPHETGGYMLECVCLGNGKGWTKCPIAEKCFDHAAGTSYVVGWETWEKPYQGMMV 180
Db 152 DTWRRPHETGGYMLECVCLGNGKGWTKCPIAEKCFDHAAGTSYVVGWETWEKPYQGMMV 211
Qy 181 DCTCLGEGSGRIITCTSRNRCNDQDTRTSYRIGDTWSKDNRGNLLQICITGNGRGWKE 240
Db 212 DCTCLGEGSGRIITCTSRNRCNDQDTRTSYRIGDTWSKDNRGNLLQICITGNGRGWKE 271
Qy 241 RHTSVQTTSSGSGPFTDVRAAV 262
Db 272 RHTSVQTTSSGSGPFTDVRAAV 293
```

Search completed: September 30, 2005, 09:21:21
Job time : 198.974 secs